Reduction, Qualia, and the Direct Introspection of Brain States

PAUL CHURCHLAND

Paul Churchland teaches philosophy at the University of California at San Diego and writes primarily in the area of philosophy of mind. He is one of the leading spokesman for physicalism, the view that mental states can be reduced to physical states—the view that Thomas Nagel and Frank Jackson attacked in the preceding selections. In this selection Churchland explains how we should understand physicalism and then why, in his opinion, certain well-known criticisms of it, including Nagel's and Jackson's, ultimately fail.

Reading Questions

1. Churchland distinguishes several different versions of the reductionist thesis. In your own words, explain the differences among them.
2. Which versions of reductionism does Churchland believe in?
3. Why does Churchland think that the anti-reductionist arguments of Nagel and Jackson fail? Do you agree?

DO THE PHENOMENOLOGICAL or qualitative features of our sensations constitute a permanent barrier to the reductive aspirations of any materialistic neuroscience? I here argue that they do not.

If we are to deal sensibly with the issues here at stake, we must approach them with a general theory of scientific reduction already in hand, a theory motivated by and adequate to the many instances and varieties of interconceptual reduction displayed elsewhere in our scientific history. With an independently based account of the nature and grounds of intertheoretic reduction, we can approach the specific case of subjective qualia, free from the myopia that results from trying to divine the proper conditions on reduction by simply staring long and hard at the problematic case at issue.

I. Intertheoretic Reduction

We may begin by remarking that the classical account of intertheoretic reduction1 now appears to be importantly mistaken, though the repairs necessary are quickly and cleanly made. Suppressing niceties, we may state the original account as follows. A new and more comprehensive theory reduces an older theory just in case the new theory, when conjoined with appropriate correspondence rules, logically entails the principles of

the older theory. (The point of the correspondence rules, or "bridge laws," is to connect the disparate ontologies of the two theories: often these are expressed as identity statements, such as \( T = m \cdot \text{inertia} \). Schematically,

\[
T_0 \supset \text{Correspondence Rules}
\]

logically entails

\[
T_n
\]

Difficulties with this view begin with the observation that most reduced theories turn out to be, strictly speaking and in a variety of respects, false. (Real gases don't really obey \( PV = \mu T \), as in classical thermodynamics; the planets don't really move in ellipses, as in Keplerian astronomy; the acceleration of falling bodies isn't really uniform, as in Galilean dynamics; etc.) If reduction is deduction, modus tollens would thus require that the premises of the new reduced theories (statistical thermodynamics in the first case, Newtonian dynamics in the second and third) be somewhat false as well, in contradiction to their assumed truth.

This complaint can be temporarily deflected by pointing out that the premises of a reduction must often include not just the new reducing theory but also some limiting assumptions or counterfactual boundary conditions (such as that the molecules of a gas enjoy only mechanical energy, or that the mass of the planets is negligible compared to the sun's, or that the distance any body falls is negligibly different from zero). Falsity in the reducing premises can thus be conceded, since it is safely confined to those limiting or counterfactual assumptions.

This defense will not deal with all cases of falsity, however, since in some cases the reduced theory is so radically false that some or all of its ontology must be rejected entirely, and the "correspondence rules" connecting that ontology to the newer ontology therefore display a problematic status. Newly conceived features cannot be identical with, nor even nomically connected with, old features, if the old features are illusory and unsubstantiated. For example, relativistic mass is not identical with Newtonian mass, nor even coextensive with it, even at low velocities. Nevertheless, the reduction of Newtonian by Einsteinian mechanics is a paradigm of a successful reduction. For a second example, neither is caloric-fluid-pressure identical with, nor even coextensive with, mean molecular kinetic energy. But an overly fluid thermodynamics (i.e., one committed to the existence of caloric) still finds a moderately impressive reduction within statistical thermodynamics. In sum, even theories with a nonexistent ontology can enjoy reduction, and this fact is problematic on the traditional account at issue.

What cases like these invite us to give up is the idea that what gets deduced in a reduction is the theory to be reduced. . . .

The point of a reduction, according to this view, is to show that the new or more comprehensive theory contains explanatory and predictive resources that parallel, to a relevant degree of exactness, the explanatory and predictive resources of the reduced theory. . . .

. . . it is to be expected that existing conceptual frameworks will eventually be reduced or displaced by new and better ones, and that in turn by frameworks better still, for who will be so brash as to assert that the feeble conceptual achievements of our adolescent species comprise an exhaustive account of anything at all? If we put aside this conceit, then the only alternatives to intertheoretic reduction are epistemic stagnation or the outright elimination of old frameworks as wholly false and illusory.

II. Theoretical Change and Perceptual Change

Esoteric properties and arcane theoretical frameworks are not the only things that occasionally enjoy intertheoretic reduction. Observable properties and common-sense conceptual frameworks can also enjoy smooth reduction. Thus, being a middle-A sound is identical with being an oscillation in air pressure at 440 Hz; being red is identical with having a certain triplet of electromagnetic reflectance efficiencies; being warm is identical with having a certain mean level of microscopically embodied energies, and so forth.

Moreover, the relevant reducing theory is capable of replacing the old framework not just in contexts of calculation and inference. It should be appreciated that the reducing theory can displace the old framework in all its observational contexts as well. Given the reality of the property identities just listed, it is quite open to us to begin framing our spontaneous perceptual reports in the language of the more sophisticated reducing theory. It is even desirable that we begin doing this, since the new vocabulary observes distinctions that are in fact within the discriminatory reach of our native perceptual systems, though those objective distinctions go unmarked and unnoticed from within the old framework. We can thus make more penetrating use of our native perceptual equipment. Such displacement is also desirable for a second reason: the greater inferential or computational power of the new conceptual framework. We can thus make better inferential use of our new perceptual judgments than we made of our old ones.

It is difficult to convey in words the vastness of such perceptual transformations and the naturalness of the new conceptual regime, once established. A nonscientific example may help to get the initial point across.

Consider the enormous increase in discriminatory skill that spans the gap between an untrained child's auditory apprehension of a symphony and the same person's apprehension of the same symphony forty years later, heard in his capacity as conductor of the orchestra performing it. What was before a barely apprehended tune is now a rich, mozaic of distinguishable elements. What was before a dimly apprehended tune is now a rationally structured sequence of distinguishable and identifiable chords supporting an appropriately related melody line. The mature musician hears an entire world of structured detail, concerning which the child is both dumb and deaf.

Other modalities provide comparable examples. Consider the practiced and chemically sophisticated wine taster, for whom the "red wine" classification used by most of us divides into a network of fifteen or twenty distinguishable elements: ethanol, glycol, fructose, sucrose, tannin, acid, carbon dioxide, and so forth, whose relative concentrations he can estimate with accuracy.

Or consider the astronomer, for whom the speckled black dome of her youth has become a visible abyss, scattering nearby planets, yellow dwarf stars, blue and red giants, distant globular clusters, and even a remote galaxy or two, all discriminable as such and locatable in three-dimensional space with her unaided (repeat: unaided) eye.

In each of these cases, what is finally mastered is a conceptual framework—whether musical, chemical, or astronomical—a framework that embodies far more wisdom about the relevant sensory domain than is immediately apparent to un tutored discrimination. Such frameworks are characterized by a cultural heritage, pieced together over many generations, and they only supplies a richness and penetration to our sensory lives that would be impossible in their absence.2

Our introspective lives are already the extensive beneficiaries of this phenomenon. The introspective discriminations we make are for the most part learned; they are acquired with practice and experience, often quite slowly. And the specific discriminations we learn to make are those it is useful for us to make. Generally, those are the discriminations that others are already making, the discriminations embodied in the psychological vocabulary of the language we learn. The conceptual framework for psychological states that is embedded in ordinary language is a modestly sophisticated theoretical achievement in its own right, and it shapes our mature introspection.
Reduction, Qualia, and the Direct Introspection of Brain States: Paul Churchland

320 PART EIGHT: UNDERSTANDING

profoundly. If it embodied substantially less wisdom in its categories and connecting generalizations, our introspective apprehension of our internal states and activities would be much diminished, though our native discriminatory mechanisms remain the same. Correlatively, if folk psychology embodied substantially more wisdom about our inner nature than it actually does, our introspective discrimination and recognition could be very much greater than it is, though our native discriminatory mechanisms remain unchanged.

This brings me to the central positive suggestion of this paper. Consider now the possibility of learning to describe, conceive, and introspectively apprehend the subtending intricacies of our inner lives within the conceptual framework of a matured neuroscience, a neuroscience that successfully reduces, either smoothly or roughly, our common-sense folk psychology. Suppose we trained our native mechanism to make a new and more detailed set of discriminations, a set that corresponded not to the primitive psychological taxonomy of ordinary language, but to some more penetrating taxonomy of states drawn from a completed neuroscience. And suppose we trained ourselves to respond to that reconfigured discriminative activity with judgments that were framed, as a matter of course, in the appropriate concepts from neuroscience.

If the examples of the symphony conductor (who can hear the Am7 chords), the oenologist (who can see and taste the glycol), and the astronomer (who can see the temperature of a blue giant star) provide a fair parallel, then the enhancement in our introspective vision could approximate a revelation. Dopamine levels in the limbic system, the spiking frequencies in specific neural pathways, resonances in the thalamic layer of the occipital cortex, inhibitory feedback to the lateral geniculate nucleus, and countless other neurophysiological niceties could be moved into the objective focus of our introspective discrimination, just as Gm7 chords and Adim chords are moved into the objective focus of a trained musician's auditory discrimination. We will of course have to learn the conceptual framework of a matured neuroscience in order to pull this off. And we will have to practice its noninferential application. But that seems a small price to pay for the quantum leap in self-appraisal.

All of this suggests that there is no problem at all in conceiving the eventual reduction of mental states and properties to neurophysiological states and properties. A matured and successful neuroscience need only include, or prove able to define, a taxonomy of kinds with a set of embedding laws that faithfully mimics the taxonomy and causal generalizations of folk psychology. Whether future neuroscientific theories will prove able to do this is a wholly empirical question, not to be settled a priori. The evidence for a positive answer is substantial and familiar, centering on the growing explanatory success of the several neurosciences.

But there is negative evidence as well: I have even urged some of myself ("Eliminative Materialism and the Propositional Attitudes," op. cit.). My negative arguments there center on the explanatory and predictive poverty of folk psychology, and their question whether it has the categorical integrity to merit the reductionist preservation of its familiar ontology. That line suggests substantial revision or outright elimination at the eventual fate of our mentalistic ontology. The qualia-based arguments of Nagel, Jackson, and Robinson, however, take a quite different line. They find no fault with folk psychology. Their concern is with the explanatory and descriptive poverty of any possible neuroscience, and their line suggests that emergence is the correct story for our mentalistic ontology. Let us now examine their arguments.

III. Thomas Nagel’s Arguments

For Thomas Nagel, it is the phenomenological features of our experiences, the properties or qualia displayed by our sensations, that constitute a problem for the reductive aspirations of any materialistic neuroscience. In his classic position paper (op. cit.) I find three distinct arguments in support of the view that such properties will never find any plausible or adequate reduction within the framework of a matured neuroscience. All three arguments are beguiling, but all three, I shall argue, are unsound.

First Argument: What makes the proposed reduction of mental phenomena different from reductions elsewhere in science, says Nagel, is that it is impossible to exclude the phenomenological features of experience from a reduction, in the same way that one excludes the phenomenal features of an ordinary substance from a physical or chemical reduction of it—namely, by explaining them as effects on the minds of human observers (437).

The reason it is impossible to exclude them, continues Nagel, is that the phenomenological features are essential to experience and to the subjective point of view. But this is not what interests me about this argument. What interests me is the claim that reductions of various substances elsewhere in science exclude the phenomenal features of the substance.

This is simply false, and the point is extremely important. The phenomenal features at issue are those such as the objective redness of an apple, the warmth of a coffee cup, and the pitch of a sound. These properties are not excluded from our reductions. Redness, an objective phenomenal property of apples, is identical with a certain wavelength triplet of electromagnetic reflectance efficiencies. Warmth, an objective phenomenal property of objects, is identical with the mean level of the objects' microscopically embodied energies. Pitch, an objective phenomenal property of sound, is identical with its oscillatory frequency. These electromagnetic and micro-mechanical properties, out there in the objective world, are genuine phenomenal properties. Despite widespread ignorance of their dynamical and microphysical details, it is these objective physical properties to which everyone's perceptual mechanisms are keyed.

The reductions whose existence Nagel denies are in fact so complete that one can already displace entirely large chunks of our common-sense vocabulary for observable properties and learn to frame one's perceptual judgments directly in terms of the reducing theory. The mean KE of the molecules in this room, for example, is currently about . . . 6.2 x 10^-21 joules. The oscillatory frequency of this sound (1 here whistled C one octave above middle C) is about 524 hz. And the three critical electromagnetic reflectance efficiencies (at 45, 52, and 63 pm) of this (white) piece of paper are all above 80 per cent. These microphysical and electromagnetic properties can be felt, heard, and seen, respectively. Our native sensory mechanisms can easily discriminate such properties, one from another, and their presence from their absence. They have been doing so for millennia. The "resolution" of these mechanisms is inadequate, of course, to reveal the microphysical details and the extended causal roles of the properties thus discriminated. But they are abundantly adequate to permit the reliable discrimination of the properties at issue.

On this view, the standard perceptual properties are not "secondary" properties at all, in the standard sense which implies that they have no real existence save inside the minds of human observers. On the contrary, they are as objective as you please, with a wide variety of objective causal properties. Moreover, it would be a mistake even to try to "kick the phenomenal properties inwards" since that would only postpone the problem of reckoning their place in nature. We would only confront them again later, as we address the place in nature of mental phenomena. And, as Nagel correctly points out, the relocation dodge is no longer open to us, once the problematic properties are already located within the mind.

Nagel concludes from this that subjective qualia are unique in being immune from the sort of reductions found elsewhere in science. I draw a very different conclusion. The objective qualia
(redness, warmth, etc.) should never have been "kicked inwards to the minds of observers" in the first place. They should be confronted squarely, and they should be reduced where they stand: outside the human observer. As we saw, this can and has in fact been done. If objective phenome-
nal properties are so treated, then subjective qualia can be confronted with parallel forthrightness, and can be reduced where they stand: inside the human observer. So far then, the external and the internal case are not different: they are parallel after all.

Second Argument: A second argument urges the point that the intrinsic character of experiences, the qualia of sensations, are essentially accessible from only a single point of view, the subjective point of view of the experiencing subject. The properties of physical brain states, by contrast, are accessible from a variety of entirely objective points of view. We cannot hope adequately to account for the former, therefore, in terms of properties appropriate to the latter domain (cf. Nagel, 442–444).

This somewhat diffuse argument appears to be an instance of the following argument:

1. The qualia of my sensations are directly known by me, by introspection, as elements of my conscious self.
2. The properties of my brain states are not directly known by me, by introspection as elements of my conscious self.
3. The qualia of my sensations ≠ the properties of my brain states.

And perhaps there is a second argument here as well, a complement to the first:

1. The properties of my brain states are known-by-the-various-external-senses, as having such-and-such physical properties.
2. The qualia of my sensations are not known-by-the-various-external-senses, as having such-and-such physical properties.
3. The qualia of my sensations ≠ the properties of my brain states.

... The fallacy committed in both cases is amply illustrated in the following parallel arguments.

1. Hitler is widely recognized as a mass murderer.
2. Adolf Schickgruber is not widely recognized as a mass murderer.
3. Hitler ≠ Adolf Schickgruber.
4. Aspirin is known by John to be a pain reliever.
5. Acetylsalicylic acid is not known by John to be a pain reliever.
6. Aspirin ≠ acetylsalicylic acid.

or, to cite an example very close to the case at issue:

1. Temperature is known by me, by tactile sensing, as a feature of material objects.
2. Mean molecular kinetic energy is not known by me, by tactile sensing, as a feature of material objects.
3. Temperature ≠ mean molecular kinetic energy.

The problem with all these arguments is that the "property" ascribed in premise 1 and withheld in premise 2 consists only in the subject item's being recognized, perceived, or known as something, under some specific description or other. Such apprehension is not a genuine feature of the item itself, fit for dividing identities, since one and the same subject may be successfully recognized under one description (e.g., "qualia of my mental state"), and yet fail to be recognized under another, equally accurate, coreferential description (e.g., "property of my brain state").

Third Argument: The last argument here is the one most widely associated with Nagel's paper. The leading example is the (quoted) character of the experiences enjoyed by an alien creature such as a bat. The claim is that, no matter how much one knew about the bat's neurophysiology and its interaction with the physical world, one could still not know, nor perhaps even imagine, what it is like to be a bat. Even total knowledge of the physical details still leaves something out. The lesson drawn is that the reductive aspirations of neurophysiology are doomed to dash themselves, unrealized, against the impenetrable keep of subjective qualia (cf. Nagel, 438 ff.).

This argument is almost identical with an argument put forward in a recent paper by Frank Jackson. Since Jackson's version deals directly with humans, I shall confront the problem as he formulates it.

IV. Jackson's Knowledge Argument

Imagine a brilliant neuroscientist named Mary, who has lived her entire life in a room that is rigorously controlled to display only various shades of black, white, and grey. She learns about the outside world by means of a black/white television monitor, and, being brilliant, she manages to transcend these obstacles. She becomes the world's greatest neuroscientist, all from within this room. In particular, she comes to know everything there is to know about the physical structure and activity of the brain and its visual system, of its actual and possible states.

But there would still be something she did not know, and could not even imagine, about the actual experiences of all the other people who live outside her black/white room, and about her possible experiences were she finally to leave her room: the nature of the experience of seeing a ripe tomato, what it is like to see red or have a sensation-of-red. Therefore, complete knowledge of the physical facts of visual perception and its related brain activity still leaves something out. Therefore, materialism cannot give an adequate reductionist account of all mental phenomena.

To give a conveniently tightened version of this argument:

1. Mary knows everything there is to know about brain states and their properties.
2. It is not the case that Mary knows everything there is to know about sensations and their properties.

Therefore, by Leibniz's law.


4. We can, I think, find at least two shortcomings in this sort of argument.

The First Shortcoming: This defect is simplicity itself. 'Knows about'... is not universal in both premises... Jackson's argument is valid only if 'knows about' is universal in both premises. But the kind of knowledge addressed in premise 1 seems pretty clearly to be different from the kind of knowledge addressed in (2). Knowledge in (1) seems to be a matter of having mastered a set of sentences or propositions, the kind one finds written in neuroscience texts, whereas knowledge in (2) seems to be a matter of having a representation of redness in some prelinguistic or sublinguistic medium of representation for sensory variables, or to be a matter of being able to make certain sensory discriminations, or something along these lines... The difference between a person who knows all about the visual cortex but has never enjoyed a sensation of red, and a person who knows no neuroscience but knows well the sensation of red, may reside not in what is respectivley known by each (brain states by the former, qualia by the latter), but rather in the different type of knowledge each has of exactly the same thing. The difference is in the manner of the knowing, not in the nature of the thing(s) known... In sum, there are pretty clearly more ways of "having knowledge" than having mastered a set of sentences. And nothing in materialism precludes this. The materialist can freely admit that one has "knowledge" of one's sensations in a way that is independent of the scientific theories one has learned. This does not mean that sensations...
are beyond the reach of physical science. It just means that the brain uses more modes and media of representation than the simple storage of sentences. And this proposition is pretty obviously true: almost certainly the brain uses a considerable variety of modes and media of representation, perhaps hundreds of them. Jackson's argument, and Nagel's, exploit this variety illegitimately: both arguments equivocate on 'knows about.'

This criticism is supported by the observation that, if Jackson's form of argument were sound, it would prove far too much. Suppose that Jackson was arguing, not against materialism, but against dualism: against the view that there exists a nonmaterial substance—call it "ectoplasm"—whose hidden constitution and nomic intricacies ground all mental phenomena. Let our cloistered Mary be an "ectoplasmologist" this time, and let her know, everything there is to know about the ectoplasmic processes underlying vision. There would still be something she did not know; what it is like to see-red. Dualism is therefore inadequate to account for all mental phenomena!

This argument is as plausible as Jackson's, and for the same reason: it exploits the same equivocation. But the truth is, such arguments show nothing, one way or the other, about how mental phenomena might be accounted for.

The Second Shortcoming: There is a further shortcoming with Jackson's argument, one of profound importance for understanding one of the most exciting consequences to be expected from a successful neuroscientific account of mind. I draw your attention to the assumption that even a utopian knowledge of neuroscience must leave Mary hopelessly in the dark about the subjective qualitative nature of sensations not-yet-enjoyed. It is true, of course, that no sentence of the form "x is a sensation-of-red" will be deducible from premises restricted to the language of neuroscience. But this is no point against the reducibility of phenomenological properties. As we saw in section 1, direct deducibility is an intolerably strong demand on reduction, and if this is all the

"sensation-of-red" to us), and see whether she can identify it correctly on introspective grounds alone, as "a spiking frequency of 90 Hz: the kind a tomato would cause." It does not seem to me to be impossible that she should succeed in this, and do so regularly on similar tests for other states, conceptualized clearly by her, but not previously enjoyed.

This may seem to some an outlandish suggestion, but the following will show that it is not. Musical chords are auditory phenomena that the young and unpracticed ear hears as undivided wholes, discriminable one from another, but without elements or internal structure. A musical chord is an integral whole, and comes to hear chords as groups of discriminable notes. If one is sufficiently practiced to have absolute pitch, one can even name the notes of an apprehended chord. And the reverse is also true: if a set of notes is specified verbally, a trained pianist or guitar player can identify the chord and recall its sound in auditory imagination. Moreover, a really skilled individual can construct, in auditory imagination, the sound of a chord he may never have heard before, and certainly does not remember. Specify for him a relatively unusual one—an F # 9th with 43th for example—and let him brood for a bit. Then play him this for four chords, one of which is the target, and see whether he can pick it out as the sound that meets the description. Skilled musicians can do this. Why is a similar skill beyond all possibility for Mary?

"Ah," it is tempting to reply, "musicians can do this only because chords are audibly structured sets of elements. Sensations of color are not."

But neither did chords seem, initially, to be structured sets of elements. They also seemed to be indistinguishable wholes. Why should it be unthinkable that sensations of color possess a comparable internal structure, unnoticed so far, but awaiting our determined and informed inspection? Jackson's argument, to be successful, must rule this possibility out, and it is difficult to see how he can do this a priori. Especially since there has recently emerged excellent empirical evidence to suggest that our sensations of color are indeed structured sets of elements.

I do not mean to suggest, of course, that there will be no limits to what Mary can imagine. Her brain is finite, and its specific anatomy will have specific limitations. For example, if a bat's brain includes computational machinery that the human brain simply lacks (which seems likely), then the subjective character of some of the bat's internal states may well be beyond human imagination. Clearly, however, the elusiveness of the bat's inner life here stems not from the metaphysical "emergence" of its internal qualia, but only from the finite capacities of our idiosyncratically human brains. Within those shrewdly structural limitations, our imaginations may soar far beyond what Jackson, Nagel, and Robinson suspect, if we possess a neuroscientific conceptual framework that is at last adequate to the intricate phenomena at issue.

I suggest then, that those of us who prize the flux and continuous of our subjective phenomenological experience need not view the advance of materialistic neuroscience with fear and foreboding. Quite the contrary. The genuine arrival of a materialist kinematics and dynamics for psychological state and cognitive processes will constitute not a gloom in which our inner life is suppressed or eclipsed, but rather a dawning, in which its marvelous intricacies are finally revealed—most notably, if we apply ourselves, in direct self-conscious introspection.

NOTES


2. The role of theory in perception, and the systematic enhancement of perception through theoretical progress, are examined at length in my Scientific Realism and the Plurality of Mind, op. cit. secs. 1–6.

3. I believe it was Paul K. Feyerabend and Richard Rorty who first identified and explored this suggestion. See Feyerabend, "Materialism and the Mind-Body Problem," Review of Metaphysics, xviii 1, 65 (September 1963); 49–66; and Rorty, "Mind, Body


Further Questions

1. Churchland thinks we should drop our familiar vocabulary for describing our experiences and start describing them in the language of neuroscience. Why? Do you agree?

2. Which is the more plausible view of consciousness—reductionism or antireductionism?

Further Readings


MARY, THE NEUROSIENTIST with only black and white visual stimuli, learns something new when she first sees a ripe tomato. All parties agree that we can describe this learned new element as “learning what the experience of seeing something red is like.” But what does this new knowledge involve?

In what is called the “knowledge argument,” anti-physicalists insist that we are dealing with new knowledge of an emergent fact, transcending all objective physical facts which can be expressed in language or which can be known from a variety of “points of view.” But physicalists, such as Nemerow [1980] and Lewis [1983], suggest that what Mary learns is not a new fact. They suggest that Mary does not learn that anything at all. Instead, Mary learns how to do something—what she gains is an ability. In having the experience of seeing something red, she learns such things as how to recognize red by looking and how to imagine the experience. Often we gain such abilities only by having the experience, though Churchland [1985] shows that there can be alternative ways to such learning.

Pointing to abilities gained shows that, although we agree that Mary has learned something new, we have an alternative to saying that Mary has learned some new (nonphysical) fact. But this defense of physicalism does not yet convince because it does not show us why what is learned seems to involve a new fact. The ability defense does not address the intuitive “subjective” or “qualitative” aspect of experience in which the nonphysical aspect of experience is supposed to lie.

What we call the subjectivity of experience is perfectly real, but, I suggest that the impression that it involves special facts arises from a confusion. Suppose we are in an advanced state of science which provides us with all the (Jackson) “physical” or (Nagel) “objective” facts. Suppose now that we arrange for a normally sighted person and a color-blind person to look at a clearly lit bright red patch. Lab assistants use microscopes and other devices to collect all the relevant physical/objective facts, which they then communicate to the two subjects. According to the knowledge argument, the color-blind subject is still missing something that the normally sighted subject has—there is something that all this factual information leaves out. The color-blind subject still does not know what it is like to see something red. All the objective facts fail to communicate the experience’s “subjective element.”

Without yet knowing what the missing element is, must we not agree that the color-blind person misses something which the normally sighted person has?

Yes, of course! When all the facts have been communicated, the color-blind person still will not have the sensation of seeing something red. He will not be in that state. No matter what information we pass on, we will not thereby produce the neurological or other physical state which constitutes having the sensation. I submit that herein lies the alleged “missing subjective el-

This essay was written expressly for this volume.