Consciousness and the Self without Reductionism:

Touching Churchland’s Nerve

**Abstract** Patricia Churchland’s *Touching a Nerve: The Self as Brain* is her most recent wide-ranging argument for mind-to-brain reductionism. It’s one of the leading anti-dualist works in neurophilosophy. It thus deserves careful attention by anti-reductionists. We survey the main arguments in this book for her thesis that the self is nothing but the brain. These arguments are based largely on the self’s dependence upon neural activities as reflected in its various impairments, its unified experiences, and its powers of agency. We show that dualism is quite compatible with this neural dependence. We argue that dualism can not only counter—but also turn the tables on—her arguments. Unlike most other critics, we focus not only on her hard problem but also her easy problems in explaining consciousness. A new non-Cartesian substance dualism is presented that avoids existing dualist causal issues. It may thus avoid perennial physicalist and dualist problems.

What am I? What is consciousness? Securing adequate answers would be tantamount to a discovery of the “holy grail” of neuroscience and philosophy. Some neuroscientists and philosophers believe that neurobiological facts alone are relevant to answering the forgoing questions. A leading example is Patricia Churchland, who has argued in her recent book, *Touching a Nerve: The Self as Brain,* that naturalism provides a better account of the self and consciousness than dualism. Churchland defines naturalism in terms of “the brain alone” (or only brain stuff) and dualism in terms of “soul stuff and brain stuff.” To clarify, we use the terms soul, self, person, and subject interchangeably. What all these terms pick out on our view is the *bearer* of rights, duties, free will, and mental properties in general, both conscious and unconscious.[[1]](#footnote-1)

Churchland is an eloquent leader for the views that comprehending minds require comprehending brains and that minds will ultimately be reduced to neuroscientific terms with no need for dualism. But we embrace dualism (and the relevant empirical data) and try to turn the tables on Churchland’s naturalism by showing that certain kinds of dualism explain minds better than neuroscience can. What’s most novel below is that while most critiques of reductionism have confronted its metaphysical problems, we focus more on its serious empirical problems in accounting for the self, its unified experience, its powers of agency, and its various impairments involving amnesia, blindsight, split brains, and paralysis. In what follows, we take a brief tour and analysis of Churchland’s naturalism and conclude that it doesn’t threaten the existence of soul stuff. Our arguments challenge a great many reductive theories of mind today.

**I. Churchland’s Brain Alone Hypothesis**

Patricia Churchland (2006, 2013) maintains that *all* mental phenomena, including consciousness and the self, point to and can be explained in terms of the brain alone: “there is just the brain—there is no separate thing, me, existing apart from my brain” (2013, p. 34). From the moral sphere to the sparks of scientific creativity, all mental phenomena boil down to our brain’s capacities: “Our moral nature is what it is because our brains are as they are; so too, for our capacities to learn, reason, invent, and do science” (Churchland 2006, p. 3). Churchland is not making the noncontroversial claim that our moral nature and mental capacities, in general, merely correlate with our brains, for any proponent of a non-materialist position—from Cartesian substance dualism (where souls and bodies are two radically different substances that can exist apart) to non-reductive physicalism—could make that claim. Rather she’s asserting that our moral nature and mental capacities in general can be explained by “the brain alone” (Churchland 2006, p. 5; see also Churchland 2013, p. 50). Churchland originally tended to deny the existence of conscious mental life, but in the more recent book we focus mainly on here she accepts its existence and tries to reduce it to brain activities (see the discussion of the hard problem below).

To bolster the preceding claim, her “brain-alone” hypothesis argues that the phenomena of neural dependence have greater consilience under this hypothesis than (Cartesian) substance dualism. Simply stated, her objection to dualism here is that the mind depends on the brain, so they can’t be radically different entities. We will now cover examples of the self’s dependence on brain activities in cases of the self’s awareness of its own body, its impairments, its decision-making, its agency, its own unity, and the unity of its experiences.

To start with, Churchland’s account of neural dependency stresses how the brain supports the self’s awareness of its own body. She argues (2013, p. 38, her italics) that the brain’s circuitry “supports a neural model of the *inner* world.” Presumably, this inner neural model maps, or spatially represents, our bodies’ skin, muscles, stomach, arms, head, et cetera. This neural model, for example, enables us to know the spatial positions of our arms and legs while biking down a challenging mountain trail or playing the drums in the dark. When this neural model breaks down, so does our capacity to represent the spatial positions of our bodily parts.

There are other data that one might enlist to show how highly dependent the mental is on the physical not only for sensory stimuli and behavior but for our conscious lives as well. This is evident in the self’s impairment due to neural damage. For example, a fall accident involving a head injury can lead to significant memory problems and personality changes. And damage to the right parietal cortex can lead to visual neglect (or a lack of awareness of the stimulus contents associated with the left half of one’s visual field) and to somatoparaphrenia (which can involve a lack of awareness of the limbs on the left side of one’s body). Further, severing the corpus callosum can lead to a split mind effect (under carefully controlled empirical constraints). Regarding the latter, Churchland maintains that split-brain results “were powerful support for the hypothesis that mental states are in fact states of the physical brain itself, not states of a nonphysical soul” (2013, p. 50). Churchland articulates her view of neural dependence deftly in the following pithy passage:

The more we know about neurology and about neuropharmacology, the more evident it is that the functions in question [reasoning, judgments, consciousness, etc.] are not remotely as independent as the classical [dualist] hypothesis asserts. On the materialist hypothesis, the observed interdependence is precisely what would be expected, but it is distinctly embarrassing to the dualist hypothesis (1996, p. 319).

Churchland also argues that self’s powers of agency depend on brain states. She argues (2006, 2013) that the brain is a “causal machine” (presumably within the logic of a linear, deterministic framework).[[2]](#footnote-2) The brain has different states at different times. A brain state at any given time is caused by antecedent conditions. Antecedent conditions include, among other events, the effects of external stimuli (e.g., encoding the prototypes of exploitation and altruism in early childhood development) and internal hormonal changes (e.g., feeling down or feeling uplifted). These sorts of antecedent conditions are both necessary and sufficient for the generation of any brain state: “If the antecedent conditions had been different, the result would have been different; if the antecedent conditions remained the same, the same result would obtain” (2006, p. 5). She motivates this claim by deploying the logic of transitivity: “If choices are brain events, and if brain events have causal antecedents, then choices have causal antecedents” (2006, p. 6). Churchland surmises that if the brain as a causal machine hypothesis were true, then the brain could be regarded as the *sole* cause of human choice and action. Hence, there would be no need to appeal to nonphysical substances or properties to explain human choice and action: “The most plausible hypothesis on the table is that the brain, and the brain alone, makes choices and decides upon actions” (2006, p. 5). An illustration about character development might be useful here. If brains cause actions, and actions cause characters, then brains cause characters. Inspired by the Scottish philosopher David Hume, Churchland maintains that if a choice could be made without reference to one’s brain (and thus without reference to the causal basis of one’s character), “the choice would be so bizarre, so ‘out of the blue’ as to raise the question of whether it was a real choice at all” (2006, p. 7; see also Churchland 2013, p. 12). In further attacking this dualist account of choices, Churchland (2013, p. 51) asks “How can energy be transferred from a completely nonphysical thing to a physical thing? Where does the soul get its oomph to have such an effect?” This is one reason that she assimilates consciousness and its causal role to neural activities.

Further elaborating on this dependence of the mental upon the neural, Churchland (2013) turns to how the self’s own unity depends on neural events. She alleges that human choice and action can be accounted for without a commander-in-chief but by mere reference to the dynamics of billions of neurons performing functions together. Self-control on her account depends solely upon the synapses between a subset of neurons in “the prefrontal cortex and subcortical structures, mainly the basal ganglia and nucleus accumbens” (2013, p. 176). From these data, she then draws the inference that there is currently no known single area of the brain (or unique module) where the will resides and carries out its functions. An analogy might be useful. Just as there is no known convergence zone for the integration of sensory information, so too there is no known single place in the brain where the self wills, conducts, or fulfills its commands.

Turning to the dependence of the self’s unified experiences on neural events, Patricia Churchland (2013) further suggests that current neurobiology is converging upon several properties central to consciousness: (1) local and global connections are required for the integration of information; (2) posterior events in the frontal areas are required for consciousness; (3) the central thalamus is required for enabling specific contents of consciousness; and (4) the common underlying mechanism of the aforementioned (1-3) neurobiological properties is the synchronous activations of widely distributed neurons: “The linkages, it is thought, may consist in synchrony in the activities of populations of neurons” (Churchland 2013, p. 147). This approach has several strengths. However, we shall argue that recent empirical evidence surrounding anesthetically induced unconsciousness substantially weakens Churchland’s purported common link here.

In these ways, Churchland argues that the self depends on neural events for its awareness of its own body, its various impairments, its decision-making, agency, its own unity, and the unity of its experiences. So, there is no need for non-neural dualist causes which radically differ from neural causes. These cases of neural dependence are better explained by her brain alone hypothesis than by dualism.

Finally, Churchland adds that her brain-only view allows for free will. Here, she embraces soft determinism’s claim that free will is self-determined instead of being externally determined or not determined at all. This is the ordinary sense of free will (2013, pp. 178ff.).

Having taken a brief tour of Churchland’s naturalism (or brain alone hypothesis), we shall now argue that dualism can not only counter—but also turn the tables on—her anti-dualist arguments concerning neural dependence, especially in regard to the self’s various impairments, unified experiences, and powers of agency. But we’ll start with the hard problem facing her reductionism, which will figure in some of our other comments below.

**II. Analysis of Churchland’s Brain Alone Hypothesis**

**Reply to Churchland’s treatment of the hard problem**

Most of Patricia Churchland’s writings address the so-called “easy problem” of consciousness concerning how brains integrate information and perform other activities associated with consciousness. Our paper is concerned mostly with the easy problems that she addresses.

But Churchland also addresses the hard problem of why these neural activities are accompanied by an experienced inner life (Chalmers, 1996, p. xi-xii). Here, Churchland (1996, pp. 395ff.) originally supported eliminative materialism, which predicts that naive folk psychology’s acceptance of consciousness will ultimately be replaced by a mature neuroscience that makes no use of consciousness. This argument has fared poorly because of Descartes’ well-known point that if we know anything exists, it’s our own consciousness. More recently, Churchland adopts the weaker contention that consciousness exists, but is reducible to neural events: “there is just the brain—there is no separate thing, me, existing apart from my brain” (2013, p. 34), so “[p]robably the soul and brain are one and the same” (2013, p. 60).

But such reductionist views face the conceivability problem, knowledge problem, and explanatory-gap problem (see, e.g., xxxx, 2016 for accounts of these). Churchland replies to critics here that our present inability to explain how her reduction works doesn’t imply that it’s false (2013, pp. 56-60). Nonetheless, it’s safe to say that until she shows how to bridge the explanatory gap between consciousness and brains, any claims of their identity will lack intelligibility. We’ll suggest a dualism below that avoids both physicalism’s problem with reductions and dualism’s problem with causality.

But, as just noted, the main focus of this paper isn’t on Churchland’s hard problem, but on her easy problems concerning what neural activities are associated with consciousness. We’ll turn to these problems now.

**Reply to Churchland’s Neural Dependence Objection**

Let’s now take a closer examination of Churchland’s brain alone hypothesis and why it doesn’t threaten the existence of soul stuff. Simply stated, Patricia Churchland’s objection to dualism here is that the mind depends on the brain, so they can’t be radically different entities.

Nonetheless, Churchland overlooks the fact here that many dualisms are compatible with this neural dependence of minds on brains. Most contemporary non-Cartesian dualists hold to a naturalistic approach (without reductionism), which could be called *naturalistic dualism*.[[3]](#footnote-3) Naturalistic dualism treats the brain as the natural seat of the soul, so the soul depends on the brain. For the naturalistic dualist, the term “soul” might only refer to (a) irreducible mental properties that are causally dependent on neural properties (e.g., weakly emergent mental properties), or to (b) an irreducible mental substance (in possession of irreducible mental properties) that strongly emerges from, and possesses causal power to act back upon its neural properties (i.e., a species of strong emergent dualism).[[4]](#footnote-4)

For option (a), John Searle’s biological naturalism will suffice for illustrative purposes, wherein specialized neural properties are thought to causally realize mental properties, such as thirst sensations (like certain molecular structures realize solidity). Those properties are reducible causally but not with respect to certain conscious phenomena, such as subjectivity, unity, and qualia (e.g., Searle 1995). The trouble here is that an explanatory gap remains. How can something entirely physical produce something nonphysical? William Hasker articulates option (b) as follows: “the human mind [or soul] is produced by the human brain and is not a separate element “added to” the brain from the outside” (1999, p. 189). Option (b) insists that the human brain is the “natural seat” of the human soul: without the human brain, there could be no human soul that undergoes phenomenal and non-phenomenal (information processing) events. Moreover, because the human brain is the “natural seat” of the human soul, the human soul derives certain properties from the human brain, such as being located in space and being functionally related to the natural world. Further, option (b) is at least open to the idea that mind could be fundamental in some sense, and thus a question about how irreducible mental entities are produced is not as mysterious as standard versions of property dualism. Hence, options (a) and (b) actually entail the neural dependence of our conscious lives on the physical.[[5]](#footnote-5) This shows that non-Cartesian brands of dualism (in which souls are irreducible to brains but don’t exist apart from brains) are not vulnerable to Churchland’s neural dependence objection.

Perhaps what Patricia Churchland has in mind when she uses the phrase “nonphysical soul” is a popular (“folk-psychological”) theistic approach to soul. However, not every theistic approach is equivalent, and some formulations entail soul-brain dependence in a very detailed way. For example, consider option (c): God simply wires the human soul (an irreducible entity that has subjective and non-subjective properties) to the human brain, so to speak; a consequence of that wiring process is that the human soul is dependent on the human brain in a very detailed way in this life. Like (b), this is strong emergence, inexplicable by physical science. Why would God make our souls highly dependent on our brains in this life? If, for example, God’s purposes were for embodied conscious agents to be highly vulnerable to each other and the environment, then God would make the soul highly dependent on the brain in this life (yyyy and cccc 2016). Now, option (c) might be an actual nonstarter for those who reject theism, including some dualists committed to option (a).  Granted.  Our point here is modest.  We only mean to suggest that option (c) illustrates how some theistically minded dualists have postulated neural dependence as a core theoretical entity. And this observation aligns with a core theoretical entity of Churchland’s hypothesis, namely neural dependence.

Another non-Cartesian dualist option with a rather different view of neural dependence is (d): the mind is seated in the brain’s electromagnetic (EM) field (xxxx & yyyy 2021).[[6]](#footnote-6) This is our own dualist field theory (DFT). It avoids reducing minds to neuroscience’s observable electrical activity by treating minds as the underlying reality of this electrical activity beyond neuroscience’s descriptions of it. We access our mental activity directly while accessing neural matter quite indirectly through instruments, reflected light, etc. Since we therefore cannot know what brain matter is really like beyond perceptions of it, for all we do know, this underlying reality could be conscious (cf. Strawson, 2006).

If all matter-energy is treated as consciousness, this yields a nonreductive monist field theory—while treating just EM activity as conscious yields DFT where EM fields are conscious and everything else is nonconscious. The brain’s sensory-processing circuitry helps generate a conscious EM field with visual images and emotions, for example. Cognitive and limbic processing yields a conscious, intelligent field—a mind. This conscious EM field is a simple, unified substance, unlike the separate neurons that generate it. For the waves comprising it reach across space as a continuous whole.

These conscious fields exert EM forces which interact with the neurons that generate them. This yields mind-brain interactions (see xxxx and yyyy (2021) for experimental evidence for this account of mental unity and causality). Here, the conscious activity has its own subjective, qualitative dynamics apart from the blind mechanics of physical science. This is evident when we weigh moral choices or even choose which foods taste best. In DFT, consciousness and its causality are thus necessarily irreducible to the physical science.[[7]](#footnote-7)

DFT’s account of neural dependence both resembles and differs from those in the other dualisms above. First, in DFT, minds depend on brains in that the brain generates the conscious EM field and helps shape its conscious content (e.g. visual images). But unlike in (a), minds act back on brains—and with dynamics irreducible to neural dynamics (yyyy and xxxx 2019; xxxx and yyyy 2021). Second, DFT also rejects (a) and (c)’s claim that consciousness emerges from brain activity that is nonconscious. Instead, the mind’s overall consciousness arises when the neural EM field joins the small-scale consciousness in the myriad neuroelectrical activities of separate cells into a simple, unified whole reaching across overall circuitries.[[8]](#footnote-8) So, while there’s no single place in brains where information processing converges (Zeki, 2003), there is a single field that unifies consciousness together. To summarize, DFT avoids some other dualism’s claims that consciousness emerges from nonconscious brains and has no causal dynamics apart from brains. DFT is emergent here in that consciousness and its dynamics aren’t explicable by natural science, as we’ll see.

DFT’s interpretation of neural dependence differs not only from that of the other dualisms above (as just explained), but also from that of Churchland. While minds do depend on brains in DFT, they do so in ontologically and causally irreducible ways. To start with, brains alone don’t account for behavior in DFT. As already noted, mental causality brings subjective, qualitative dynamics to behavior. In contrast, Churchland’s reductionism doesn’t explain how neuronal circuits alone can have subjective, qualitative dynamics (e.g., choosing which food tastes best or which face looks prettier). Also, DFT rejects her claim that brains alone exist. This reductive (or eliminative) claim leaves her unable to explain the conscious percepts, thoughts, and feelings we obviously possess. As already noted, DFT also explains how minds get their unity, while Churchland offers no explanations here. All this arguably turns the tables on her claim that experimental evidence about brains has greater consilience under her brain alone hypothesis than under dualism.

In sum, all forms of naturalistic dualism hold that the human soul is highly dependent on the human brain in very detailed ways. Unlike some others, option (d) adds that consciousness doesn’t emerge from brains that lack consciousness—and that the soul has its own causal dynamics irreducible to the brain’s dynamics. A broader theoretical implication of all the naturalistic dualisms above is that we would expect to discover highly detailed physical correlates upon which our phenomenal and non-phenomenal lives depend. Not surprisingly, some dualists these days are, in fact, interested in discovering those highly detailed physical correlates for the sake of developing a testable theory.[[9]](#footnote-9) Thus, whatever Churchland might mean by a “nonphysical soul,” one thing is clear about the issue of neural dependence: over the past 2400 years very few, if any, noteworthy dualists have denied that the “soul” causally depends in a very detailed way on the physical. This dualism is, of course, compatible with souls having their own causal dynamics and interacting with brains (yyyy 2013; yyyy and xxxx 2019; xxxx and yyyy 2021).

**Neural Dependence: Hard versus Easy Cases**

Churchland might push back on grounds that, so far, our response to the neural dependence objection to dualism has not addressed any of the hard cases, such as those involving impairments to the self, such as amnesia and blindsight. One might claim that these cases reveal key properties of consciousness that are dependent upon and probably explainable in terms of our neural hardware without remainder. We think these key properties of consciousness boil down to *easy* problems of consciousness (which are really just problems about physical structures and functions associated with consciousness). Philosophers point out here that the former concerns hard problems (about experience) but the latter concerns easy problems (about functions and their related structures). For example, *amnesia* concerns problems about information encoding, storage, and access, but not the harder problem of subjective experience. When a computer begins to lose its capacity to encode, store, and access information, we do not thereby conclude that it has begun to lose its subjective experience. Presumably, it never had subjective experience. So why would we ever conclude that persons are subjects of experience on the basis of encoding, storing, and accessing information? Perhaps we would not, unless we’ve already conflated the hard problem with one or more of the easy problems of consciousness. Also, none of the specific data surrounding amnesia cases entail the absence of a subjective point of view. It’s just that one’s experience of being a subject is truncated in time because the physical structures responsible for encoding, storing, and accessing information have undergone damage. One could still know what it is like to be a subject of experience, if only over brief periods of time (yyyy & cccc).

In the case of blindsight, one could draw a distinction between *information processing* versus *phenomenal experience* (see Flannigan 1992; Holt 2003). Information processing refers to a system’s capacity to encode and react to information in the natural world. Information processing, however, does not require phenomenal experience. Something as simple as a motion detector on a security system engages in information processing. Our brains, too, have evolved several specialized subsystems that engage in information processing of the stimuli they distinguish, including the capacity to encode motion (e.g., in visual area 5, or V5). Yet a subject cannot be phenomenally aware of motion (or of other stimuli) without information processing. The relation is asymmetric: phenomenal experience requires information processing, but the reverse does not hold (Holt 2003; see also Chalmers 1995; xxxx 2008; 2013b; xxxx and cccc 2016). This suggests that while blindsight patients are still engaged in information processing of certain stimuli; they’re just not phenomenally aware of the stimuli they process and differentiate (see also pppp, yyyy, bbbb 2018).

While blindsight suggests that brains can fully explain some cognitive responses, there is currently no evidence that brains alone can explain higher cognition, including creative imagination.

In any case, the data surrounding neural dependence map elegantly onto dualist options (b), (c), and (d) above: human subjects undergo both phenomenal and non-phenomenal events; thus some of what we do (as embodied subjects) is likely influenced by and dependent on the non-phenomenal (information processing) events that we undergo via our brains. Contrary to Churchland’s inference above, the data surrounding neural dependence do not have greater consilience under the brain alone hypothesis.

**Reply to Churchland’s Split-Brain Objection**

Patricia Churchland could push back in another way. She could suggest that split-brain phenomena run counter to a subject (or self or soul) that’s irreducible to the brain. If the brain can split, but the soul cannot, then don’t split brains challenge the existence of soul stuff? In fact, Churchland maintains that split-brain results “were powerful support for the hypothesis that mental states are in fact states of the physical brain itself, not states of a nonphysical soul” (2013, p. 50). However, our DFT in fact doesn’t assume the soul is indivisible, for severing connections between the hemispheres blocks the flow of electrical currents and their highly localized fields between hemispheres. This can yield two separate subjects in DFT—which aligns with behavioral evidence (such as the battling of opposing hands) in these patients. So split brains don’t undermine DFT but help illuminate its explanatory power as well.

**Reply to Churchland’s Purported Common link**

Further, recall Patricia Churchland’s claim that the common link to, or neural correlate of, consciousness is the synchronous activations of widely distributed neurons: “The linkages, it is thought, may consist in synchrony in the activities of populations of neurons” (Churchland 2013, p. 147). However, extensive data show that neuronal synchronization is present (and actually strengthens) in the relevant striate and extra-striate regions of the primary sensory networks when an animal is rendered unconscious by means of anesthesia (Bola et al 2018; Hudetz 2006, 2009; Imas et al. 2005; Mashour 2013a, 2013b). Furthermore, posterior-to-anterior processing of the primary sensory networks is actually preserved during anesthetic-induced unconsciousness in humans (Ku et al., 2011; Mashour 2013a, 2013b) and non-human animals (Imas et al. 2005). Not only do these data punch squarely against Churchland’s purported common link, but they also challenge recent related hypotheses grounded in the claim that neuronal synchrony plus amplification (via attention and/or some other mechanism) could count as the distinctive correlate of consciousness. For example, recent advocates of this claim include Carruthers (2015), Dehaene (2014), Levy (2014) and Prinz (2012). By implication, this widely purported common link (neuronal synchrony) cannot be the unique (or distinctive) correlate of consciousness (yyyy 2019).

While several more objections to her view of common links exist (see yyyy 2006, 2007, 2010, 2019; yyyy and xxxx 2019; xxxx and yyyy 2021), we’ll focus on a foundational one that directly challenges the connectionist assumption of her hypothesis. Churchland’s common link to consciousness assumes that spatial connections (synapses) and temporal connections (synchronous activations) exist between all relevant neuronal subassemblies. However, current data suggest otherwise: the properties of an object are not only distributed in space, owing to a lack of direct anatomical links between processing sites[[10]](#footnote-10), but they are “distributed in time” as well, owing to a lack of temporal connections between processing sites (Zeki 2003, p. 215; Zeki 2015b; yyyy 2007; 2010; 2019; yyyy et al 2020). For example, while a subject is exposed to an object, Zeki and colleagues have systematically observed a temporal gap as great as 80 milliseconds between visual processing subsystems, V4 and V5. In effect, the current evidence reveals an *asynchronous* relation between *geographically separate* processing sites—an observable datum that cuts directly against Churchland’s (and many other scholars’) purported common link to, or neural correlate of, consciousness.[[11]](#footnote-11)

Here, again, DFT may turn the tables on Churchland’s claim that experimental evidence has greater consilience under her brain alone hypothesis than under dualism. For DFT explains how perceptions and other kinds of conscious activities are unified, while (as just explained) her synchronic approach doesn’t. In DFT, the brain’s single, continuous EM field in diffuse currents can unify numerous dispersed circuits to produce a single unified consciousness, even where these circuits don’t synapse with each other. The neural EM field is thus a better candidate for the neural correlate of consciousness (i.e., unified consciousness) than her candidate of neural synchrony. Experimental evidence for this conclusion comes from Koch et al. (2016). They argue that locally activated EEGs track conscious perceptions across brains better than other events such as synchrony. This EEG evidence correlates unified perceptions with local neuroelectrical fields.

**Reply to Churchland’s Naturalist Account of Agency**

As we observed earlier, Churchland also maintains that if a choice could be made without reference to one’s brain (and thus without reference to the causal basis of one’s character), “the choice would be so bizarre, so ‘out of the blue’ as to raise the question of whether it was a real choice at all” (2006, p. 7; see also Churchland 2013, p. 12). In transitive form, for example, if brains cause actions, and actions cause characters, then brains cause characters. We think her linear, deterministic approach to agency raises deeper puzzles in clinical neuropsychiatry about a subject’s (or self’s) capacity to rewire its brain (or to engage in subject-directed neuroplasticity), say, in the case of stroke, OCD or certain other pathologies. We do not think the brain alone hypothesis can adequately accommodate such feats, though the brain is most certainly involved in them. According to our (strong) emergent subject hypothesis (ESH), when a brain’s EM field is activated, an ontologically irreducible (and simple) subject emerges and possesses causal power (in its own right) to influence the structures and functions of its brain. It’s important to note that our rejection above of weakly emergent consciousness doesn’t conflict with our acceptance here of a strongly emergent subject. The former tries to explain the rise of the conscious subject and its causality in terms of physical science (as in the first kind of naturalistic dualism discussed earlier). The latter doesn’t try to explain the conscious subject and its causality in terms of physical science, instead this subject intelligibly arises with its own dynamics from simpler forms of consciousness in sensory, memory, and [?] emotional neural activities (yyyy and xxxx 2019; xxxx and yyyy 2021; yyyy 2010, 2013a, 2019; yyyy et al 2020). ESH is a species of non-Cartesian substance dualism. Is there empirical evidence to support ESH with respect to mental causation, thereby challenging Churchland’s claim that exceptions to the brain alone hypothesis are bizarre? There is growing evidence along these lines from recent neuroplasticity studies.

Consider the key role that modified constraint-induced movement therapy (CIMT) can play in the aftermath of severe cerebral strokes. This species of therapy involves constraining the intact arm so that the patient has to use the paretic arm. When this occurs, a self is carrying out an operation that requires considerable mental effort (yyyy et al 2020). Abundant evidence suggests that patients who engage in such cognitive (or subject-directed) behavioral therapy are actually stronger after the stroke event (Yadav et al 2016). For example, severe cardiac strokes can cause significant neurological damage in a part of one of the motor cortices. If severe damage occurs in the right motor cortex, immobility on the left side of the body usually follows. But immobility of this sort is not necessarily permanent. Some stroke patients can, through sustained conscious effort, rewire adjacent neural areas in the same motor cortex, enabling them to execute (overt) actions prior to the stroke event, say, moving one’s left arm. In most stroke cases, the conscious subject usually causes adjacent neural wiring in the motor cortex impacted by the stroke. However, there are rare cases of stroke patients with severe right motor cortical damage who have been known to recruit and rewire neural areas in their left motor cortex. When this unusual form of neuroplasticity occurs, the left motor cortex is used not only for right-sided movements but for left-sided movements as well. The initial stages of this unusual form of subject-directed neuroplasticity can create a strange phenomenal experience for the subject. For example, a patient described the phenomenology of the initial stages of the rewiring process as involving a strange battle between her left and right thumbs. While trying to exercise her left thumb, the patient would experience her right thumb moving along with her left thumb in an almost mirror-like fashion. She experienced her consciously directed left thumb movements in tandem with her non-consciously directed right thumb movements. That strange experience occurred because, in the early stages of the subject-directed neural rewiring process, her conscious effort to bring about left thumb movements had recruited some neurons in the left motor cortex that were normally used to carry out right thumb movements. Through sustained conscious effort, she eventually carved out new neural pathways in her left motor cortex for left-sided movements that did not involve right-sided interference. Through sheer will power this patient totally defied her doctors’ prognosis; she was told that she would never be able to move her left side, owing to the magnitude of neurological damage in her right motor cortex. These empirical data underscore the power of subject-directed neuroplasticity and thus how “plastic or adaptable” the brain can be in relation to the conscious voluntary power of a subject (Schwartz and Gladding 2011, pp. 37-38; yyyy 2013a). This subject-directed plasticity counters Churchland’s brain-alone hypothesis. Our view here [is] a sensible account of this phenomenon, rather than being bizarre, as she claims.

More recently, yyyy et al (2020) advocate a strong emergence hypothesis and argue on logical and empirical grounds that when a subject (or self) is engaged in CIMT, and other tasks that require considerable mental effort, the subject is causing its brain to change (i.e., subject-directed neuroplasticity). Subject-directed neuroplasticity has also been deployed to effectively treat obsessive-compulsive disorders, addiction disorders, attention disorders, and to preempt the onset of full-blown seizures (yyyy et al 2020; yyyy 2013a, 2013b).

The philosophical upshot is that we possess causal power to reorganize our neural wetware. Having this causal power not only increases the survival value of subjects in possession of it, but the associated empirical evidence above runs directly counter to Churchland’s deterministic, ‘brain alone’ hypothesis. By implication, our emergent subject hypothesis provides a better fit of this data than materialism in general and Patricia Churchland’s ‘brain alone’ hypothesis in particular. We conclude that dualism is not floundering but quite feasible compared to Churchland’s naturalism today. The evidence of subject-directed neuroplasticity works hand in hand with DFT and ESH to counter Churchland’s brain alone hypothesis and its implication that mental causation is merely neural causation.

DFT can go further and turn the tables here on Churchland‘s brain-alone hypothesis in two ways. First, her anti-dualist brain-only hypothesis relies partly on the argument that souls lack causal powers (which she calls “oompf”) to affect brains (2013, p. 51). Yet DFT counters this argument by showing how souls can be the underlying nature of the neural EM field which has considerable causal powers in brains. Second, her hypothesis fails to establish that only brains exist—after all, we exist. As argued above, we know our own consciousness exists, and it’s not clear how consciousness and its causality can be nothing but brain activity as her hypothesis claims. In contrast, DFT readily explains consciousness and its causality.

**Implications for Free Will**

Churchland’s naturalism (only brain stuff) raises issues surrounding the free will of selves (which is another example of the self’s powers of agency). Our field theories may be able to avoid such issues. We begin with a brief description of three standard theories of free will.  In hard determinism, our choices are determined, so we can’t choose otherwise than we actually did, as free will requires. In indeterminism (or libertarianism), our choices aren’t determined (at least not by physical causes), so we have free will. We’re free to act as we choose (which involves liberty), and we do so without determinism (which requires indeterminism). In soft determinism, our choices are both determined and free, for freedom is self-determinism (autonomy). Only in this view are free will and determinism compatible.

Patricia Churchland sees indeterminism’s claim that free will requires indeterminism as being irrelevant to practical life. She embraces soft determinism’s claim that free will is self-determined instead of being externally determined or not determined at all. This is the ordinary sense of free will, which is used in law courts and addresses different sorts of diminished self-control (2013, pp. 178ff.).

But this free will may be illusory, given her naturalist view that selves (or minds) are just brains.  This naturalist view arguably leaves us as mere puppets of the neurobiological forces controlling brains. We don’t control these or any other laws of nature, so we don’t really control ourselves—we’re just puppets mistakenly thinking we’re free. Further, if we’re determined by neurobiological forces, then arguably we aren’t free to do otherwise than we do—we’re locked into what we do and can’t be held responsible for doing otherwise. Churchland doesn’t reply to these standard anti-free-will arguments.

Our DFT can also adopt soft determinism, yet in ways that may be able to avoid the preceding issues.  To start with, we base autonomy on top-down causation involving emergent causality based around qualia that transcends physical science. We thus avoid Churchland’s causal reductionism, which arguably renders us into puppets of neurobiological forces. Furthermore, following Kane’s (1996) ideas on free will, we combine elements of soft determinism and indeterminism, (which involves altering standard soft determinism). We allow for indeterminism in top-down causation at quantum levels, which avoids deterministic claims that we couldn’t have acted otherwise (see Xxxx & Yyyyy, 2021, on all this).  Note that this indeterminism doesn’t make our choices erratic, it just helps randomly create options for our more deliberate levels of thinking to draw upon. In these ways, we arguably avoid Churchland’s standard free will problems.

Moreover, in our DFT, emotions and other qualia play a vital role in soft determinism’s core idea of autonomy. But Churchland fails to account for qualia in terms of neural information processing. She thus fails to account for the crucial role of emotions in free will. What is this role of emotions? To start with, our choices involve weighing our inner feelings in ways that transcend the principles of neural networks. We construct our own autonomous goals in the form of ideals and values imbued with strong feelings. Through these feelings, the will draws its controlling power from the primal emotional powers it controls, much like power steering (Lorenz, 1966, 1977). These passionate principles bring their own unfolding logics to our creative imaginations, a logic of ideas. New ideologies, art forms, etc. arise with lives of their own. They harness the external forces that partly create them, transforming them into new directions. This is the vital role of our qualia’s inner life in free will. Here, our wills are autonomous, not puppets of neurobiological forces. Without our qualia, we’d be vegetative robots, devoid of feelings and purposes. This is, in effect, what we become under Churchland’s brain-alone hypothesis.

**Works Cited**

Bola, M., Barrett, A., Pigorini, A., Nobili, L., Seth, A., Marchewka, A. (2018), Loss of consciousness is related to hyper-correlated gamma-band activity in anesthetized macaques and sleeping humans, *NeuroImage* 167: 130-142.

Buszaki, G. (2006) *Rhythms of the brain*. Oxford: Oxford University Press.

Chalmers, D. (1995) Facing up to the problem of consciousness. *Journal of consciousness studies* 2 (3): 200-219.

Chalmers, D. (1996) *The Conscious Mind*. Oxford: Oxford University Press.

Carruthers, P. (2015). *The Centered Mind*. Oxford: OUP.

Churchland, Patricia (1996) *Neurophilosophy*. Cambridge: MIT Press. Copyright 1986.

Churchland, Patricia (2013), *Touching a Nerve: The Self as Brain* (W. W. Norton & Company, New York).

Churchland, Patricia (2006), Moral decision-making and the brain (pp. 3-16). In Judy Illes (editor) *Neuroethics*. Oxford: OUP.

Dehaene, S. (2014). *Consciousness and the brain.* New York: Penguin Group.

Flannigan, O. (1992). *Consciousness Rediscovered*. Cambridge: MA: The MIT Press.

Holt, J. (2003). Blindsight and the nature of consciousness. New York: Broadview Press.

Hasker, William (1999). *The Emergent Self.*  Ithaca: Cornell University Press.

Hudetz, A. (2006). Suppressing Consciousness: Mechanisms of General Anesthesia. *Seminars in Anesthesia* 25: 196-204.

Hudetz, A. (2009). Feedback suppression in anesthesia: Is it reversible? *Consciousness and Cognition* 18: 1079-1081.

Levy, N. (2014). *Consciousness and Moral Responsibility*. Oxford: OUP.

Lowe, E. J. (2008) *Personal Agency*. Oxford: OUP.

Ku, S. et al (2001). Perfrontal inhibition of frontal-to-parietal feedback connectivity is a neurophysiologic correlate of general anesthesia in surgical patients. *PLoS ONE* 6: e25155.

Mashour, G. (2013a). Cognitive unbinding. *Neuroscience and Behavioral Reviews* 37: 2751-2759.

Mashour, G. (2013b). Consciousness, Anesthesia, and Neural Synchrony. *Anesthesiology* 119: 7-9.

Kane, R. (1996) *The Significance of Free Will*. Oxford: OUP.

Lorenz, K. (1966) *On Aggression*. Methuen.

Lorenz, K. (1977) *Behind the Mirror*. Harcourt Brace Jovanovich.

Prinz, J. (2012). *The Conscious Brain*. Oxford: OUP.

Schwartz, J., Gladding. R. (2011). *You are not your brain*. New York: Penguin Group.

Strawson, G. (2006), ‘Realistic monism’, *Journal of Consciousness Studies,* 13 (10-11), pp. 3-31.

Yadav, R. Sharma, R. Borah, D. Kothan, S. (2016). Efficacy of modified constraint induced movement therapy in the treatment of hemiparetic upper limb in stroke patients: a randomized controlled trial. *Journal of Clinical Diagnostic Research*. 11: YC01-YC05. doi: [10.7860/JCDR/2016/23468.8899](https://dx.doi.org/10.7860%2FJCDR%2F2016%2F23468.8899)

Zeki, Semir (2003). The disunity of consciousness*. Trends in Cognitive Sciences* 7 (5): 214-218.

1. This isn’t meant as a hard and fast point—for example, apes arguably have self-awareness but aren’t persons. [↑](#footnote-ref-1)
2. This view contrasts with the logic of nonlinear dynamical systems, in which the mind is thought to act back upon the brain through recurrent interactions, also known as downward causation by mind on brain. [↑](#footnote-ref-2)
3. This phrase was introduced by David Chalmers (1995) to name his preferred brand of dualism. I use it here because it clearly applies (in different senses) to other contemporary brands of dualism. [↑](#footnote-ref-3)
4. Option (a) traces all the way back to Aristotle (see yyyy 2008), while option (b) is advocated by several contemporary philosophers (e.g., see Hasker 1999; Lowe 2008; yyyy 2013a). [↑](#footnote-ref-4)
5. Moreover, option (b) possesses the explanatory power needed to address Jaegwon Kim’s famous pairing relations objection to Cartesian dualism. Also, entity dualism entails property dualism, but the reverse does not follow. [↑](#footnote-ref-5)
6. Field theories of mind have been held by renowned thinkers like Kohler, Libet, Eccles, and Popper (see xxxx, 2013, for references). They draw on considerable experimental evidence, withstand past criticisms, and help avoid neuroscience’s issues in explaining the mind’s consciousness, unity, and causality. [↑](#footnote-ref-6)
7. This DFT is a modified form of Lowe’s non-Cartesian substance dualism (NCSD) (xxxx & yyyy, 2021). DFT addresses two issues Lowe wasn’t clear about. First, Lowe treated subjects as unified, simple substances but wasn’t clear how subjects come to be this way. Second, Lowe insisted that subjects lack forces and energy, but wasn’t clear how the conscious purposes of subjects affect bodily behavior without any forces or energy transfers. Treating subjects as the underlying reality of unified, simple neural EM energy fields thus helps clarify Lowe’s NCSD. [↑](#footnote-ref-7)
8. In DFT, consciousness doesn’t mysteriously pop into existence from what lacks consciousness, as in some other dualisms. Instead, consciousness is the underlying reality of fundamental electromagnetic energy fields (beyond what we observe of them via electrodes and other instruments). Yet only in brains is consciousness organized to form images, thoughts, and overall minds. [↑](#footnote-ref-8)
9. For example, see yyyy 2007, 2013a, 2019; yyyy and xxxx 2019; yyyy et al 2020; xxxx and yyyy 2021. [↑](#footnote-ref-9)
10. While there are no direct anatomical links between brain areas, as connectionism assumes, there are often as explained below diffuse ion currents (with their strong internal fields) among these circuitries (in, for example, cortical modules), as our field approach assumes. [↑](#footnote-ref-10)
11. It might be argued that while no single frequency of synchronized activity links geographically separate processing sites, various frequencies couple together to do so, as when theta (or alpha and beta) frequencies help control gamma frequency activities in memory, attention, and working memory (Buszaki, 2006). But such “cross-frequency coupling” just serves to reinforce the neural EM fields that do the actual binding (Jones, 2019). Indeed, it’s hard to see how synchronic or coherent EM frequencies alone could ever unify consciousness. One of several reasons is that, given the sheer numbers of brains on earth, it’s highly likely their EM frequencies will often cohere—yet there’s no evidence that they actually share any unified consciousness. [↑](#footnote-ref-11)