POST-PRINT AUTHOR COPY
ACCEPTED IN *ORGANIZATIONAL RESEARCH METHODS* ON JANUARY 26, 2017

PHILOSOPHICAL FOUNDATIONS OF NEUROSCIENCE IN ORGANIZATIONAL RESEARCH: FUNCTIONAL AND NONFUNCTIONAL APPROACHES

Richard P. Bagozzi[a]

University of Michigan

Nick Lee

Warwick Business School

Acknowledgements: The authors wish to express special gratitude to the Associate Editor and Reviewers for their many constructive suggestions and recommendations on an earlier draft of this article

[a]The order of names is alphabetical. Both authors contributed equally to this paper.

Abstract

Neuroscience offers a unique opportunity to elucidate the role of mental phenomena, including consciousness. However the place of such phenomena in explanations of human behavior is controversial. For example, consciousness has been construed in varied and conflicting forms, making it difficult to represent it in meaningful ways without committing researchers to one species of consciousness or another, with vastly different implications for hypothesis development, methods of study, and interpretation of findings. We explore the conceptual foundations of different explications of consciousness and consider alternative ways for studying its role in research. In the end, although no approach is flawless or dominates all others in every way, we are convinced that any viable approach must take into account, if not privilege, the self in the sense of representing the subjective, first-person process of self as observer and knower of one's own actions and history, and the feelings and meanings attached to these. The most promising frameworks in this regard are likely to be some variant of non-reductive monism, or perhaps a kind of naturalistic dualism that remains yet to be developed coherently.

Key words: consciousness, neuroscience, organizational behavior, dualism, monism, reductionism, mental events, qualia, folk psychology

What is mind? No matter. What is matter? Never mind. (Homer Simpson, quoting George Berkeley).

"It is certain that this me, that is my soul, by which I am what I am, is entirely and truly distinct from my body". (Descartes, Meditations Metaphysiques, p. 222).

"...you, your joys and sorrows, your memories and ambitions, your sense of personal identity and free will, are in fact no more than the behavior of a vast assembly of nerve cells and their associated molecules". (Crick, 1994, p. 3).

We ourselves are large-scale, complex instances of something both objectively physical from outside and subjectively mental from inside. (Nagel, 2012, p. 42).

In recent times, there has been a veritable explosion of interest in applying neuroscientific insights and methodological approaches to questions related to social science research (e.g. Bagozzi et al., 2013; Salvador and Folger, 2009; Waldman, Balthazard, and Peterson, 2011). Advances in neuroscience technologies have also been accompanied by significant methodological debates regarding the limits of inference of neuroscience research (e.g. Kosslyn, 1999; for a recent debate, see Kievit et al., 2011a,b; Bagozzi, 2011a, Barrett, 2011). Further, our increasing knowledge of the brain has moved a number of more fundamental questions about the nature of individual subjective experiences, consciousness, morality, value, free will, and indeed just what it means to be human, from the domain of abstract philosophy into more mainstream scientific discourse. These questions have been explored in depth by many neuroscientists and philosophers (e.g. Block, 2007a; Burge, 2007; Crick, 1994; Damasio, 1994, 2010; Dennett, 2006; Nagel, 2012; Tellis, 2011)

Within the organizational and management science literature, a number of scholars have provided general discussions of the role of the human brain and how neuroscience might benefit the science and practice of management (e.g. Ashkanasy, Becker, & Waldman, 2014; Becker, Cropanzano, & Sanfey; 2011). Such work has on occasion approached issues regarding the possible limits of inference from neuroscience data (i.e., what it can tell us about physical brain activity itself, see Lee, Senior, & Butler, 2012; Senior, Lee, & Butler, 2011). Further, Healey and Hodgkinson (2014) provide a useful critical realist framework for neuroscience in management and organizational research, which views cognition as embodied and socially situated, and neuroscience as one building block for knowledge development. This notable work aside, however, there is a conspicuous lack of attention paid in the management/organizational research literature to many of the important philosophical issues raised above – that is, what does brain activity *mean*? Whereas commentaries have appeared on occasion critiquing the ability of neuroscience to explain complex organizational phenomena (e.g. Lindebaum, 2015; Lindebaum & Zundel, 2013), the reductionist ideas criticized in such work merely sketch but one of the significant philosophical implications of neuroscientific advances. However, a coherent understanding of the conceptual underpinnings of research into mental phenomena and consciousness by use of neuroscience methods is of fundamental importance to the future development of theoretically and practically useful organizational research. We herein substantively deepen the current level of philosophical discussion in organizational and management science, and give guidance for empirical research in this growing area.

Consider the following questions. If some alien was able to make a perfect copy of you, right down to the very last subatomic particle, would that copy-you have the same mental experience or consciousness as the real-you? In essence, would it "be" you? Similarly, how would the "transporter" in *Star Trek* work? If it disassembled your physical body and brain, and

perfectly reassembled them somewhere else, how would it transport your consciousness? Would "you" still be "you"? Or, imagine you were dying, and technology allowed you to "port" your consciousness into an android, or even into a totally virtual world inside a computer. How would you do this? What, *exactly*, would you "port"? Next, consider what evidence you have that anyone other than yourself even *has* consciousness. What do you observe other than other people's physical features and behaviors? Does this make them "conscious"? Are you *sure*? Is your dog conscious? How does that square with your answers to the earlier questions? Finally, imagine we were able to create a perfect replica of a physical human brain, housed inside a perfect replica of a human body. Would this simulacrum be conscious, experiencing the world subjectively, like you? Or would it be a non-conscious "zombie"? These are classical thought experiments that provide an entry point into thinking about the nature of our concept of the mind, and its relationship to the physical world, including our brain (e.g., Dennett, 2013).

There are of course various ways of addressing such issues, the main points of which we will briefly introduce readers to in due course. One thing however is certain, these questions are not simply interesting ways of passing the time. The question of whether the mind and brain are the same thing is nothing less than the defining question of our time for the future of organizational (and by extension, social) science. First, the growing body of researchers interested in applying neuroscientific research methods to organizational research questions must address the question of exactly what purpose such research serves (Healey & Hodgkinson, 2014). For example, if it can be shown that physical brain events map directly to mental events, which actually exist and can have causal impacts on other mental and physical events or properties, then one must wonder what purpose neuroscientific approaches to organizational research serve. This is because in a social context such as organizational life, it could be argued that it is far more important to understand and explore higher-level psychological, social, or organizational theory

concepts as the primary drivers of important organizational consequences, than physical brain activity. This is seemingly at the root of criticism of neuroscientific organizational research by authors such as Lindebaum and Zundel (2013). But, if it can be shown that considering mental events as "real" is questionable, then this throws into doubt virtually the entirety of the last century of organizational research. In such a case, almost none of the concepts explored in existing research would be justifiably "real", and as such it is hard to see how it can be reasonable that such concepts can have causal effects on the world (Kim, 1993). This in turn has impacts on how we can have any valid knowledge of the organizational world *without* incorporating knowledge of the physical brain's role. At best, such concepts as "satisfaction", "charisma", "leadership", "intention", and "emotion" (among the seemingly thousands of others) could be thought of as metaphors, possibly connected to nothing real at all (and thus unmeasurable by their nature), which throws into doubt the idea that we have any knowledge about organizations at all. Or as philosopher Fodor (1990, p. 156) says; "if…[the causal power of mental concepts is not]…literally true, then practically everything I believe about anything is false and it's the end of the world".

By ignoring the question of the relationship between mind and body, or worse being unaware of it, organizational research risks becoming at best a sterile irrelevance, and at worst misleading or harmful. Take for example the now-infamous 2011 *New York Times* op-ed written by Martin Lindstrom, entitled "*You Love Your iPhone, Literally"*. This article drew inference from brain imaging studies to suggest that individuals feel the same emotion towards their mobile phones and other devices as they feel towards humans that they love. This claim exemplifies the naïve understanding of exactly what neuroscientific data can tell us about complex psychological and social phenomena, and was devastatingly debunked by a letter to the *NYT* by Russell Poldrack and 44 other neuroscientists a week later. Whereas this example is from the popular

press, peer-reviewed organizational / management science literature is also far from perfect in this regard.

Of course, whereas many management scholars would likely benefit from enhanced technical knowledge of neuroscientific methods, we maintain that it is greater understanding of the philosophical implications of neuroscientific research that would be most beneficial at this time. As such, it is disconcerting that so little consideration has heretofore been paid to the issues that we raise in this article. Perhaps organizational scholars are in agreement with Stephen Hawking, who states outright that "philosophy is dead. Philosophy has not kept up with modern developments in science…Scientists have become the bearers of the torch of discovery in our quest for knowledge" (Hawking & Mlodinow, 2010, p. 5; c.f., Silbersweig, 2015). Yet this seems a strangely hubristic position to take, particularly given how often cherished scientific theories are supplanted. Rather, we concur with Thomas Nagel:

> *"One of the legitimate tasks of philosophy is to investigate the limits of even the best developed and most successful forms of contemporary scientific knowledge…intellectual humility requires that we resist the temptation to assume that tools of the kind we have now are in principle sufficient to understand the universe as a whole. Pointing out their limits is a philosophical task, whoever engages with it, rather than part of the internal pursuit of science…[it is a task of]…trying to recognize what can and cannot in principle be understood by certain existing methods" (Nagel, 2012, p. 3-4).*

It is our view that, as yet, organizational scholarship has not considered such issues in enough depth, and it is this that we intend to remedy herein. Specifically, we begin by outlining the main philosophical lines of argument regarding the relationship of the mind to the brain. Subsequently, we develop a multi-level framework for organizational research that allows us to take account of evidence for both mental events, and physical brain events, and how they can be integrated into a

holistic research methodology. Of course, we do not claim to solve the mind-brain problem, which may very well be intractable. However, the approach we outline here allows researchers to move forward in a more coherent and informed fashion, and clearly sets out a place for neuroscience in social science research, and perhaps more importantly a plausible place for organizational research itself as a valid avenue for knowledge development in the contemporary world.

PHILOSOPHICAL FOUNDATIONS: PHYSICALISM, MIND, AND BRAIN

In order to begin to address the questions of exactly what neuroscientific research can tell us about human, social, and organizational phenomena, we must first address a critical philosophical question. Specifically, is there anything more to the mind than its material – i.e., those nerve cells and molecules cited by Crick (1994)? In order to address the above question, we must introduce a number of critical philosophical concepts, which frame the different positions on this issue, beginning with *physicalism* (with respect to the terms materialism and physicalism, we will regard them here as synonyms; see Stoljar [2015] for a history and comparison of the two concepts.). For purposes of discussion, we can identify three approaches to thinking about consciousness which contrast purely physical with purely mental, and with mixtures of the two approaches (see Figure 1). One, dualism, has perhaps the oldest history and is the most intuitive, being especially considered in Enlightenment times due to Descartes. Descartes advocated substance dualism, the theory that consciousness and the brain are two different kinds of things. The brain is physical, consciousness is in some sense non-physical, even supernatural in some accounts. Property dualism, as proposed in one variant by Nagel[1] (2002, 2012), posits that people observe physical reality objectively from the outside, and mental reality subjectively from the inside; here consciousness and the brain are still separate substances, but both are in some sense natural, yet consciousness is neither physical nor supernatural. Chalmers (2010) advocates a

somewhat similar view, termed naturalistic dualism, where there are objective physical entities and subjective non-physical entities useful for understanding our minds, and both are natural.

[Figure 1 about here]

Dualism in one sense sets the tone for the other two perspectives on consciousness shown in Figure 1. Idealism claims that all things are ideas, and the mental is the foundation for reality (Guyer & Horstmann, 2015), whereas substance monism asserts that all things are physical (Schaffer, 2016). It is difficult at this time to envision how idealism might inform empirical scientific discourse on neuroscience applied to social science research, so we will not have any more to say about it here.

Substance monism has evolved in two directions: reductive physicalism and non-reductive physicalism. Both variants maintain that all things are physical. But reductive physicalism sees conscious properties as physical properties, whereas non-reductive physicalism allows that (some) conscious properties cannot be completely reduced to the physical. The most developed instances of substance monism in philosophy, and the most frequently followed forms of substance monism guiding neuroscience, are elaborations on themes of functionalism. The general form of functionalism for conscious states, events, or processes is a causal one, where consciousness gets its meaning from the role it plays in a causal system with (physical) inputs and outputs, and perhaps other conscious states, events, or processes (e.g., Levin, 2013; see also Block, 2007 a, b). More specifically, a hypothesized conscious state, event, or process plays an intermediary role between a causal input, such as found in experimental manipulations, and its relationship to causal outputs, such as behavioral performances. The intermediary role may be represented as a construct or ideal, sometimes referred to as a hypothetical construct. Intermediary roles can be complex, such as found in cognitive or emotional processes (for an

example of such a process in an experiment where three cognitive processes interact in a 2x3x2 design, see Figure 1c, Bagozzi, 2011a; Tam et al., 2010).

More specifically, under reductive functionalism, the meaning of a conscious construct is achieved by its dependence on a physical manipulation and the role it has in producing a physical output, where the conscious construct operates on the information in the manipulation to specify a functional property. Under non-reductive functionalism, the meaning of a conscious construct is attained by its role connecting a causal input to a causal output, and where the conscious construct relates to a functional property in a way weaker than identity. We will give specific elaborations of functional accounts of consciousness in the latter half of this article.

Other forms of substance monism should be briefly mentioned for completeness (see Figure 1). According to the identity thesis, mind and brain are identical (Smart, 2014). Eliminativism claims that all or some supposedly conscious states do not exist (Ramsey, 2013). Emergentism maintains that consciousness appears or originates somehow out of physical brain processes, but is not reducible to them (O'Connor & Want, 2015). Finally, anomalous monism asserts that consciousness must be physical because psychological laws cannot exist in the same sense as physical laws, yet consciousness can cause and be caused by physical events (Yalowitz, 2014). Unlike reductive and non-reductive functionalism, these species of substance monism are difficult to apply in neuroscience approaches and will not be considered further herein.

The philosophical variants of mind are obviously numerous and complicated to both discuss and comprehend. Each is also controversial and subject to multiple objections and defenses (see entries and references in the Appendix). To date, there seems to be little consensus on this topic. As a consequence, we have chosen to discuss what neuroscience research can inform us concerning social science research by focusing on the approaches that dominate inquiry in cognitive neuroscience at the present: namely, physicalism of both reductive and nonreductive

variants, as expressed in functionalism, and close alternatives. But first let us explore a way of framing some broad issues as a prelude to our consideration of more specific methodological and empirical issues in the latter half of the article.

Figure 2 presents a very simplified conceptual schematic of what we term a "naïve model" of the relationships between our brain, our mind, and our behavior. Specifically, we have a physical brain, which somehow plays some role in our experiencing of our thoughts and feelings, which play a role in our behavior (the paths represented by arrows a and c). At the same time, most people are comfortable with thinking that there are some behaviors that are what they might term "automatic" or "instinctual" (e.g., recoiling from a threat), which do not require conscious experiencing to result in behavior (path e; see e.g., Le Doux, 2015). Similarly, typically, many people think that the environment we live in can produce some of our mental experiences, which are processed by the brain (paths d and b), but, and at the same time, some aspects of the environment can be processed without conscious experience – usually termed, "subconsciously" (path f). In this naïve model, the mind (i.e., an umbrella term relating to our subjective conscious experiencing of the world around us, and our thoughts and feelings, some of which become habitual and automatic) plays some mediating role between our physical brains and our functional behaviors, some of the time.

[Figure 2 about here.]

Obviously, the naïve model in Figure 2 appears to contradict a substance monist view, being as it posits some non-physical entity of "mind", which plays an active causal role in our behaviors. However, substance monism holds that everything that exists is physical, and that our mental experiences and consciousness are, subject to certain caveats, identical to their corresponding brain materials and interactions. As such to maintain a substance monist view, we must eliminate or in some way reduce everything that is above the dashed line which demarks the

physical from the non-physical, or else argue that seemingly subjective experience *supervenes* on the material (e.g., Kim, 2000, 2005; see also entry in Appendix, and Figure 6b). Supervenience means in essence that our mental properties are wholly dependent on our physical properties, whereas not being exactly the same thing. However, even if this is the case, Kim (1993) claims to show that it is impossible for mental events (i.e. the mind) to have unique causal powers over and above physical events (i.e., the brain). Thus, we are left with our brain, our behavior, and the links between them (paths e and f). Of course, as our understanding of the universe has evolved, our understanding of exactly what counts as physical matter has also changed, but the key point regarding substance monism is that the entirety of that which exists is in some sense physical. Currently, we have no satisfactory way of conceptualizing our subjective mental experience as physical, so from this point of view we must conclude that it does not exist in the way it naïvely seems to exist to us (Figure 2).

Substance monism thus seems to jar with our intuitions, and our actual experiences, of being human, and as such considerable time has been spent in exploring how it comes to be that we seem to experience the world through what feels to us like a first-person subjective point of view. In other words, why does it "feel" like we are more than just brains responding to the environment? Just what actually is our conscious experience? Why does it feel like anything at all? Consider again the earlier example of a perfect simulacrum of a human, which is known as a *philosophical zombie* (or p-zombie). According to what we know about the physical brain, a p-zombie would be functionally and biologically indistinguishable from a normal human. The problem therefore being that, if such a creature is logically possible, then a) how do we know anyone other than ourselves is conscious at all, and as such, b) how can subjective consciousness be defined in a physicalist way?

The world of subjective experience, or "feeling", is what philosophers term the *phenomenological* world, with actual subjective conscious experiences called *qualia*. Whereas numerous empirical scientists and philosophers have tried to reconcile the substance monist view with our subjective phenomenological experience in various ways (e.g. Pinker, 2015; Wegner, 2002), one well-received and reasonably archetypal model has been provided by Dennett (e.g. Dennett, 1991; Dennett, 2003). It is impossible here to explain Dennett's model in any brief way that does it justice, but in essence, his theory is that subjective consciousness we each experience is the result of multiple parallel brain activity responses to the myriad stimuli that make up the world, which create "Multiple Drafts" of perception, some of which are then promoted to functional roles by the brain, which acts in this case like a virtual von Neumann machine (i.e., a serial information processing architecture using stored programs, but which can only process one instruction at a time). The "mind" is thus a collection of simplified evolved modules (e.g., ducking, avoiding predators), which partly act in fulfilling their hard-wired purpose, but are also on occasion enlisted to perform other functions. We present commentary in a concluding section to this article, on a similar approach to consciousness emerging in the neuroscience literature.

Models such as Dennett's provide plausible explanations of how it comes to be that a purely physical neurobiological system could come to have what *seems like* a subjective consciousness. Whereas it is clearly substance monist in nature, Dennett (1991, see p. 459) resists the urge to classify his theory as solely fitting into one of the specific reductive categories introduced above (identity theory, eliminativism, reductive functionalism), suggesting that to do so is broadly meaningless, because it shares aspects of each. Nevertheless, the clear goal of Dennett's and other such models is one of psychophysical reductionism, the specific claim that eventually consciousness, and thus by extension the mind itself, will be completely explainable by physics and chemistry (Nagel, 2012). In doing so, models such as Dennett's redefine our

subjective phenomenological experience of consciousness to be at best some kind of useful fiction, a deception, misconception, or hallucination created by our brains as a result of their evolution, or at worst an unintended side effect of the complex system that is our brain. In essence, our subjective experience is considered from this point of view to be an illusion (Pinker, 2015). After all, if mental experience is not physical, it cannot be real in the sense of the naïve model in Figure 2. Whereas this may be unsatisfying from the perspective of our actual experience of being human (i.e., we *feel* like there must be more to it than that), this type of model has the significant advantage of placing the study of the brain and mental experience into the realms of our current physicalist models of science.

However, the advantages of psychophysical reductionism come at a heavy cost, which we must be willing to bear if we are to commit to this view. As alluded to above, this cost is seemingly nothing less than the removal of the concept of consciousness in scientific discourse as we actually experience it. Perhaps because of this, the claim of Dennett and others to have "explained" consciousness in this way has been challenged from different angles (see Dennett, 2003). First, Pinker (2015), when discussing his own computational model of the mind (which shares many features with Dennett's), claims outright that the question of how subjective experience and consciousness (which he terms *sentience*) emerge out of information processing "beats the heck out of me! I have…no idea of how to begin to look for a defensible answer, and neither does anybody else" (Pinker, 2015, p.146). He goes on to say, "but saying that we have no scientific explanation of sentience is not the same as saying that sentience does not exist at all. I am as certain that I am sentient as I am certain of *anything*." (p. 148). This is what we might term the *brute fact* view of consciousness; that it is there, but we have no idea how or why, nor any idea of where to start looking (see also Fodor, 1990; Wright, 2000). Pinker claims that this does not impair our understanding of how the mind works (as does Wegner, 2002). Even if this is true

though, it seems to be an unsatisfying position to be in. Dennett (2003) suggests that both Pinker and Wright have misunderstood his theory as claiming consciousness itself does not exist. Be that as it may, at best Dennett appears to *redefine* consciousness away from our subjective mental phenomenological experience, claiming the identity of "conscious [experience are], information-bearing events in the brain – since that's all that's going on" (Dennett, 1991, p. 459), and dedicates an entire chapter to dismissing qualia, claiming that they are "treacherous pitfalls" (p. 433) and explicitly that "there are no such properties as qualia" (p. 455). We consider qualia later in the article.

Nagel (2002; 2012) provides similar criticisms to those of Pinker above, but rather than trying to defend conscious experience as simply a brute fact, he approaches the mind-body problem through the lens of analytic philosophy. This effort draws from Kripke (1980), which includes the statement, "Materialism, I think, must hold that a physical description of the world is a *complete* description of it, that any mental facts are 'ontologically dependent' on physical facts in the straightforward sense of following from them by *necessity*" (p. 155, emphasis added).

So, a physicalist (i.e. substance monist), reductive functionalist account of consciousness (see Figure 1), such as Dennett's, must in the end be eliminative to some extent, redefining our subjective experience of consciousness as an illusory outcome of other brain functions (e.g., information processing), concluding that mental events *are* physical events in the brain and nothing more. Thus, with scientific advances, we will be able to discover the necessary connections between mental and physical events, in the same way that science allowed us to understand once-mysterious things such as the molecular composition of water, and fire as oxidization. However, this is might be characterized as problematic, because consciousness does not seem to be *constituted* by physical brain states (i.e., necessary), but instead to be somehow *produced* by it – which is a contingent not a necessary relationship, implying that mental states

and brain states cannot be identical (Nagel, 2012). Gray sums this up in his explanation of why the problem of explaining consciousness is so difficult: "nothing that we know so far about behavior, physiology, the evolution of either behavior or physiology, or the possibilities of constructing automata to carry out complex forms of behavior is such that the hypothesis of consciousness would arise if it did not occur in addition as a datum of our own experience; nor, having arisen, does it provide a useful explanation of the phenomena observed in those domains" (Gray, 1995, p. 660). As such, we are in a quandary at this point if we wish to retain a substance monist view. Despite its attractions, Dennett's model does not really solve this problem, being as it at most provides a plausible mechanism by which purely physical brain activity *could* result in the experience of a (perhaps illusory) subjective consciousness, but not how consciousness is the *necessary* consequence of such activity. Feyerabend (1963) "solves" this problem before it begins, by suggesting mental events do not exist at all, and Pinker (2015) solves it by treating consciousness as a simple brute fact, which we may not ever be able to grasp, and which has no importance in terms of how the mind works. Neither approach is attractive, with one dismissing a fundamental part of our actual experience of the world as illusory, and the other suggesting it is essentially unknowable.

Nagel (2002, 2012) approaches this issue by speculating that the only logical way forwards is to consider neither the physical nor mental alone as an adequate representation of reality, and instead to look for a third concept, which contains both physical and mental as necessary constitutive components (rather than simply correlated or causally / contingently related). The characteristics of such a concept appear to currently be out of reach, but Nagel (2002) goes into significant depth to explain exactly what such a concept would need to do in order to succeed. Again, we cannot do justice to Nagel's (2002; 2012) theory in the short space we have available, but in essence it would have to be a new theoretical construction, referring to

something real, which necessarily explained both the subjective phenomenological experience (inside) and the physiological manifestations (outside) of some given experiential episode. For example, wanting to work harder on the job so as to attain an objective, or being pleased with one's performance on the job, each have a unique mental phenomenological concept that we experience, but also a physiological concept which describes the physical state of the brain (located in the reward system of the brain and related to the role of dopamine in the former case, and residing in hedonic "hot spots" in the brain in the latter case and involving opioids and endocannabinoides, where the physical processes compromising both produce such phenomenological experiences as pleasure e.g., Berridge, 2009). The correct way to think of this, in Nagel's view, is to "think that these two ways of referring – by the phenomenological concept and the physiological concept – pick out a single referent, in each case rigidly, but that the logical link cannot be discovered by inspecting the concepts directly: Rather, it goes only through their common necessary link to the referent itself" (Nagel, 2002, p. 44). It is important here to be clear that this is not a model where the unspecified new concept *causes* the mental and physical. Rather, that the relationship is *constitutive*, in that the new concept is such that its being a mental state and a brain state are due to the very nature of the concept itself, in that it could be nothing else (i.e., it is a necessary truth). Such a theory would also have to explain the connections we see between manifestations of relevant phenomena as necessary characteristics of the new concept, as well as satisfy what is called *upward entailment*, in that all the manifest properties of our aforementioned experiential episode are completely explained (entailed) by our new concept/theory. Only if we do this are we able to conclusively make the claim that brain activity without consciousness is impossible or incomplete, because by their definition they would be linked necessarily, not contingently, and a p-zombie would be logically impossible. Such a new theory would also have to be compatible with the purely physical explanations of the physical

components of the same event (in other words it should not contradict what we already know about the purely physical aspects of the world).

Even leaving aside the positing of a partly non-physical entity as Nagel does, this might seem a challenging set of criteria, even impossible. However, it should be clear that scientific progress has regularly furnished us with theories of exactly this character regarding the physical world, which took manifestly contingent and unexplained correlations (such as we currently observe between mind and brain), and explained them with a deeper theory and set of concepts which accounted for all of the higher-level behaviors. The theory of fire as oxidation, heat as molecular movement, and electromagnetic field theory, all convert observable (yet at the time mysterious) contingent correlations into necessary consequences of the nature of their newly-proposed phenomena, converting previously separate phenomena into a single explanation that explains both, and providing us with testable hypotheses. For example, "the very same atomic (or molecular) agitation that accounts for increased pressure against the walls of the container accounts for increased temperature of the gas within" (Nagel, 2002, p. 49), which explains the macro-level positive correlation between the observable changes in temperature and pressure of a constant volume of gas as necessary due to the underlying property of molecular movement. So, a theory of the characteristics stipulated by Nagel is not a scientific impossibility, in fact such theories have been posited regularly throughout history. What about the conjecture of some non-physical concept? Well, whereas Nagel does not provide any real speculations on the nature of such a concept, is it too far-fetched for us to consider even the possibility of such a thing? Perhaps it is, but given a model of physics which has come up with stranger and stranger purely physical explanations, Nagel (2012) suggests perhaps there are other ways of getting to the truth other than the substance monist. Nagel himself denies that this is a substance dualist approach *a la* Descartes, which he regards as untenable. However, as explained earlier, Nagel's approach

seems more akin to a property or natural dualism, rather than the neutral monism that he occasionally alludes to, hence our characterization of it in Figure 1. In this sense, he argues that his approach is reductive in nature, but not physicalist, that consciousness can be reduced, but not purely to physical matter.

[Figure 3 about here]

Figure 3 provides a necessarily stylized and simplistic take on both positions that have been discussed, using the same concepts as Figure 2's model for ease of understanding. Whereas we do not claim these representations to be perfect, we think they provide a useful comparative take on the issues, and should help readers grasp the concepts of importance. We take a reasonably agnostic view at this point on which of these basic positions best represent the truth of the matter, but nevertheless, one or the other seemingly *must* be true. Either the brain is identical to the mind, or it is not. They are the universe of possibilities, necessarily. Of course, Dennett's and Nagel's approaches are far from the only specific ways of theorizing about these two possibilities. However the implications of all approaches to each of the conjectures are basically the same. As was shown above, both lead to some discomfiting conclusions, and neither is without cost in terms of implications for organizational (and general social science) research. On the one hand, unless organizational science wishes to become in essence a branch of functional neuroscience, a place needs to be found for our subjective experiences and higher-level phenomena of organizations and groups (Healey and Hodgkinson, 2014). This is at the roots of criticisms such as those of Lindebaum and Zundel (2013). Yet, discomfort with the implications of reductionism alone is not enough to discount the possibility that indeed, it may be the truth. If we wish to maintain a valid place for organizational research in science, we must provide some reasonable defense for the position that mental concepts are valid explanatory mechanisms in organizational theories. Indeed, specific variants of this general challenge face all social science,

whether we like it or not. The last half of this article explores some options aimed at pointing to a reconciliation or convergence in points of view.

We believe that management / organizational (and to a large extent social / psychological in general) researchers have in the main not realized that taking a position on these issues has important implications for doing valid research, just as understanding the technology and experimental paradigms of neuroscience does. Yet where there are many articles on the latter in the organizational literature, there are virtually none on the former. Certainly, it seems to us that substance monism is the (perhaps unspoken) position assumed by many organizational / management scholars to underlie the neuroscientific approach (although see our later discussion under limitations where the critical management perspective is considered). That said, it could be that practicing organizational researchers place this issue in the "too hard" basket and carry on doing what they do as a pragmatic matter, which to date has been to accept neuroscience's reliance on reductionism. However, any approach, pragmatic or otherwise, will use concepts whose conceptual/logical foundation is in need of analysis, and likely falls under one or more of the positions regarding the philosophy of consciousness and mental events noted earlier in the article.

Below, we elaborate on the specific issues of importance for neuroscientific organizational research which we believe are considerably under-appreciated at present, and present an attempt to derive a logically and conceptually coherent framework for future empirical work in the area. In particular, we make a tentative attempt to use an essentially functionalist methodological framework (with certain modifications or deviations also considered) to approach, or at least not contradict, the operationalization of Nagel's conceptualization of the mind-body problem, which does not reject the possible reality of subjective mental experiences. That said, Nagel's approach is too underdeveloped to receive explicit operational consideration in

our presentation. As such, in a more pragmatic sense, we show how researchers can employ existing neuroscientific, behavioral, and social-psychological methods in parallel to generate more valid insights into important organizational research problems.

FUNCTIONALISM AND THE NATURE OF MENTAL STATES (AND CONSCIOUSNESS)

We have sketched above the main principles of the various approaches to theorizing about the nature of subjective mental phenomena like consciousness. However, because the space available to our presentation does not allow us to elaborate further on the many philosophical doctrines regarding the nature of mental states, we have chosen to use functionalism as an organizing framework to speak about a number of contentious issues. Whereas functionalism has primarily become associated with reductionist explanations, functionalism itself in different forms can underpin both reductive and non-reductive philosophical perspectives on consciousness. We also consider approaches akin to functionalism but deviating in interesting ways.

In organizational research that deals with consciousness or mental phenomena, it is helpful to consider different levels of analysis. The most basic level, potentially applicable to organizational research, concerns chemical, physical, or molecular explanations of organizational behavior; above this level are cognitive, emotional, motivational, or volitional variables involved in such explanations as social identity, attachment styles, emotional engagement, inspirational leadership, and intuitionist ethics in organizations; and at a higher level yet are such larger scale phenomena as cooperation, power, and other organizational interpersonal or small group processes, and macro organizational phenomena. Under reductionism, the middle and higher levels are in some sense identical, reducible, or partially reducible to the most basic level, or perhaps alternatively might be thought to emerge from them.

Assuming that organizational research desires to interpret or represent the nature of consciousness or mental variables, rather than eliminating them altogether or assuming them away, one or more of the many varieties of functionalism provide a way of doing this. Indeed, most research in neuroscience and psychology, upon which organizational researchers are beginning to draw upon, follows explicitly or implicitly some variant of functionalism. Formally, Figure 4a presents a simple schematic of a basic functionalist model. (note that one deviation from classic functionalism, which treats the mediating variables or processes as hypothetical, is the alternative where such variables or processes are measured; see Figure 6 and accompanying discussion below).

[Figure 4 about here]

Importantly, functionalism is agnostic in terms of whether mental states exist, are real, or are physical or nonphysical (Block, 2007b). To take a general example of functionalism in organization research, consider Figure 4b. Here a visual stimulus manipulation is administered (e.g., photos of emotionally-charged scenes versus a control group stimulus lacking emotional content); this serves as input to a cognitive event that is hypothesized in some way to transform exposure to the stimulus; and a causal output, say in the form of a behavioral reaction or action, is measured and predicted to occur according to a theory linking causal input, to cognitive response, to causal output. In a neuroscience context, the general form here is expressed in Figure 4c, which is a baseline for elaborations developed hereafter. Perhaps the most useful elaborations on functional approaches are the following. Instead of limiting functional investigations to single levels of inquiry, such as predictions of brain states, it is informative to combine levels of analysis within any particular study. For example, the study of consciousness or mental states can be enhanced by employing both brain state measurements and subjective measurements as causal outputs, where both of these are linked explicitly according to theory (or theories) to the proposed

consciousness or mental state. Cacioppo and Decety (2011, p. 169) describe the general logic of one form of multilevel research as follows: "breaking down the component of interest and showing how, based on the prior literature on the brain, different predictions about what circuits should be activated can be derived from two or more theories".

An example of one variant of multilevel research is shown in Figure 5a. Here causes of a hypothesized mental event (or process) are manipulated along with a control condition, and two physical, behavioral outputs are registered. Notice the two outputs are shown correlated (see r and curved line connecting them). The brain state(s) and measured subjective state response(s) are parallel outcomes of the hypothesized effects of the mental event. Each outcome could be predicted by different theoretical processes, where the processes might be governed by different, but also possibly overlapping, mechanisms. Depending on the level of explanation and representation of the mental event and its predictions, the linkages from mental event to the outcomes could consist minimally of two different (but possibly overlapping) brain circuits, one brain circuit and one folk psychology mental event, or two folk psychology mental events (see discussion of folk psychology below). The larger the number of physical, behavioral, and mental outputs measured and the greater the correlations among them, the greater the behavioral evidence for the conceptual foundation underlying the theory of the hypothesized mental event. This is analogous to the notion of convergent validity as operationalized in construct validity studies. Moreover, by including one or more additional behavioral measures of outputs, where these are expected to have no or at least weaker theoretical connections to the mental level under scrutiny, we would obtain further evidence of validity, in this case discriminant validity (see Bagozzi, 2011b, for a discussion of construct validity, how to test for it, and related issues).

[Figure 5 about here]

There are different kinds of theoretical specifications linking hypothesized consciousness or mental states or processes to behavioral outputs. One example is shown in Figure 5b. Here researchers were studying theory of mind processes (i.e., the ability to infer the mental states or volitions of other people) in sales account managers, which they manipulated through one experimental and two control groups (where the former was a number of audio narratives of theory of mind processes transpiring between salesperson and customer, and the latter were narratives for a control with no interpersonal mentalizing processes and a control with unlinked sentences, see Dietvorst et al., 2009). Two classes or levels of behavioral outcomes were measured: one consisted of brain states measured by fMRI and found in previous research to capture theory of mind processes (medial prefrontal cortex, temporal poles, temporo-parietal junction activation); the other was a 4-dimensional, 13-item psychological scale derived to measure subjective report estimates of theory of mind (ToM) processes. The overall ToM scale correlated r = .69 with the activation of the right, and r = .61 with the activation of the left, temporo-parietal junctions, and r = .69 with the activation of the medial prefrontal cortex levels of activation. The correlations between the activation of the temporal poles and ToM scale did not reach statistical significance (not shown in Figure 5b for simplicity). Figure 5b shows a case of functionalism where hypothesized conscious states are related to outcomes directly linked to these specific conscious states, according to theory and past research.

Another kind of theoretical specification linking hypothesized consciousness or mental states or processes to behavioral outputs is shown in Figure 5c. Here researchers were studying empathy processes in sales account managers, which they manipulated with video clips of people expressing positive and negative emotions, and neutral faces and moving geometrical shapes were used as two control conditions (Bagozzi et al, 2012). Two classes or levels of behavioral outcomes were measured: one was comprised of brain states (measured by fMRI procedures) that

previous research showed indicated activation of mirror neurons (supplemental motor area, precentral gyrus, post central gyrus, pars opercularis/pars triangularis), executive functions (precuneus R, precuneus L), perception (inferior parietal lobule, superior parietal lobule), facial recognition (fusiform gyrus), and a 5-item scale, measuring subjective report estimates of customer orientation. Unlike the ToM study described above, where subjective reports as output behaviors directly matched the mental state of theory of mind (Figure 5b), in this study empathy was the mental state to be captured by a functionalist argument, but the subjective behavioral output was an expression or outcome of empathy, not direct measures of empathy, designed to capture behavioral orientations toward others based on empathy (in the form of customer orientation). The customer orientation scale correlated r = .55 with the supplemental motor area, r = .72 with the precentral gyrus, r = .72 with the post central gyrus, r = .67 with the pars opercularis/pars triangularis, r = .57 with the right precuneus, r = .57 with the left precunles, r = .54 with the inferior parietal lobule, r = .68 with the superior parietal lobule, and r = .57 with the fusiform gyrus (where, by referring to these regions in the brain, we mean their activation compared to a control group(s)).

A third kind of theoretical specification linking hypothesized consciousness or mental states or processes to behavioral outputs is shown in Figure 5d. Here researchers were studying Machiavellianism in sales account managers. Machiavellianism is a mental, social condition characterized by the manipulation of others for personal gain. It seemingly shows some similarity to psychopathy, which Baron-Cohen (2011) asserts reflects a deficit in empathy. The study in question manipulated two distinct facets of empathy: taking the perspective of others, a cognitive state, and feeling the emotional state of others, an affective state (Bagozzi et al., 2013). The authors manipulated ToM and (affective) empathy in two separate studies (see Figure 5d). Using brain state predictions similar to those described above for ToM (Dietvorst et al., 2009) and

empathy (Bagozzi et al., 2012), the authors found that Machiavellianism is associated with depressed ToM processes in the form of reduced perspective taking, and elevated empathy in the form of automatic (not conscious) resonance with the feeling of others. The correlations of subjective responses on a Machiavellianism scale with the respective brain states summarized in Figure 5d confirmed these results. Thus two commonly thought dimensions of empathy, which have been presumed to go together and even define empathy (based on self-report research in the literature) were, in fact, found to be disassociated and relate in opposite ways, when brain states, measured with fMRI, and subjective reports, measured with an established scale, were employed. Here two mental states thought to underpin Machiavellianism (a cognitive state of perspective taking based on ToM predictions; and an affective state of empathy based on psychological theory) achieve meaning through their transformative role between the causal manipulations and dual causal outputs (i.e., multiple brain state measurements and subjective report measures of Machiavellianism, where each class of output was linked via two theories and prior research to the proposed cognitive and affective mental events).

MULTILEVEL NEUROSCIENCE RESEARCH AND INTERNAL AND EXTERNAL VALIDITY

Multilevel neuroscience research such as illustrated above serves to deepen understanding by addressing both internal and external aspects of validity of inquiry. We can think of multilevel research as grounding neuroscience in physical and subjective phenomena. Three principles proposed by Cacioppo and Bernston (1992; see also Cacioppo & Decety, 2009) apply when behaviors across levels of mental organization are investigated. The principle of *multiple determinism* maintains that any behavior at one level of organization can have multiple antecedents within or across levels of organization. This implies that any behavioral phenomenon at one level of organization can function to explain another variable at the same or different

levels. The principle of *nonadditive determinism* claims that the whole is not necessarily predicted by the simple sum of its parts. For example, a psychological response may not be decomposable and explicated by specification of other neurological processes by themselves, but only in conjunction with differences in behavior at an interpersonal, group, or other social level (see Cacioppo & Bernston, 1992, p. 1024, for a specific case). Finally, the principle of *reciprocal determinism* asserts that mutual influence can occur between biological and social variables to determine behavior. This should not be interpreted necessarily as simultaneity or mutual causation within a single point in time, we would assert, but rather typically occurs recursively, such as might happen when social variables influence psychological variables through their effects on certain neural processes. Other consequences and directions of effects might transpire, such as reflected in so-called downward causation or emergent phenomena. For a multilevel example where social and biological levels of analysis complement each other, see Cacioppo et al. (2000).

A purely substance monist account is not compatible with the issues outlined above (e.g. Kim, 1993). Yet by crossing levels of analysis as described above, and in accordance with the three grounding principles, one obtains deeper means into internal validity and richer extensions into external validity of mental phenomena than can be normally obtained with single-level research. Such an approach to research provides a window into consciousness or mental events not easily viewed through the narrow prisms of the individual philosophical points of view on consciousness noted in Figure 1 by themselves. Rather, by combining physical behaviors (e.g., brain states or activations) with measured subjective experiences and crossing levels of analysis, as guided by the grounding principles mentioned above, we perhaps take a small step in the direction sketched by Nagel's (2012) proposal that (albeit never specified in detail) consciousness is something more than physical matter and in some way requires, what he terms,

both an ahistorical constitutive account grounded in physical phenomena (either reductive or emergent) and a historical account rooted in intentionality and evolutionary principles.

Organizational research resides squarely in the dialectical tension between basic research and applied research. We have emphasized more the basic or scientific side of organizational neuroscience in this article, although we believe the multilevel approach we have outlined bridges basic research and applied research through concerns with internal and external validities, which attempt to connect consciousness with behaviors, theory, and subjective experiences. But being also a professional field, organizational neuroscience, we would argue, should go beyond internal and external validity, as discussed above, to consider more explicitly ecological validity. One way to do this has been practiced within each of the three studies mentioned in Figures 5b-5d. That is, implications of the mental states that were studied with measurements of brain states and subjective experiences, within an experimental functionalist paradigm, were then further tested in the field with survey methods on new samples of managers (see Dietvorst et al., 2009; Bagozzi et al., 2012; Bagozzi et al., 2013). Going further, the comparatively simple task of incorporating higher-level organizational and social phenomena in the framework would allow (at least in our view) an operationalization of the critical realist framework put forward by Healey and Hodgkinson (2014). Below, we provide three other ways to add to or qualify multilevel research, and challenge pure functionalism.

SUPERVENIENCE, EMERGENCE, AND FOLK PSYCHOLOGY

Beyond functionalism, the nature of consciousness and mental events can be explored through three other philosophical lenses, each of which has appeared briefly above: supervenience, emergence, and folk psychology. To explicate these concepts in a way that maintains continuity with the discussion so far, and simplify our presentation, we will use

functionalism as a framework for illustration, although the terms in question could be discussed in other ways.

Kim (2005, p. 33) defines mind-body supervenience as follows (see also McLaughlin and Bennett, (2014), for a classical definition and history of the concept):

> *"Mental properties strongly supervene on physical/biological properties. That is, if any system s instantiates a mental property M at t, there necessarily exists a physical property P such that s instantiates P at t, and necessarily anything instantiating P at any time instantiates M at that time."*

Supervenience thus expresses or interprets the relationship between properties of mental events and physical properties. The relationship is often taken to be a kind of (necessary) dependence, not causation, because causality is maintained typically to occur between physical events or properties, not between a physical and (nonphysical) mental event or property (Kim, 2005; see Bagozzi, 2011b, pp. 93-95).

Unlike an identity or reductive relationship between mental events or properties and physical events or properties, which are taken to be identical with each other, a supervenience relationship maintains the mental event or property gets its meaning in a dependent sense from a corresponding physical event or property, and therefore the relationship is nonreductive. Figure 6 shows a supervenient relationship and contrasts it with functional and causal relationships. Notice first in Figure 6a, a hypothetical mental event gets its meaning through the functional relationships we discussed above. The relation between the measured brain state and a subjective mental event, where both these are presumed to be translated by the hypothetical mental event in response to a causal event (e.g., an experimental manipulation) to produce the two causal (in this case measured) outputs, is due to the functional connections (i.e., the role played by the hypothetical mental event). The relationship between the brain state and subjective mental event

is a statistical inferred correlation, r. The relationship could be represented as a function in a regression equation, but because the brain state and mental event are measured, perhaps at the same point in time, not manipulated, we could not interpret this strictly as a causal relationship, hence reference to r.

[Figure 6 about here]

Contrast the functional case to the supervenient case in Figure 6b. Here the physical manipulation causes a change in a brain state (which can be measured, say, by fMRI), and the relation of the brain state to a hypothetical mental event is one of supervenience. If we have two physical events and two corresponding hypothetical mental events, the relations might be depicted as in Figure 7 (see Kim, 2005, p. 45). Notice here that P_1, and P_2 are linked causally, and M_1 and M_2 get their meaning from $P_1$and P_2, respectively, but no causal relation exists (under supervenience) between M_1 and M_2.

[Figure 7 about here]

Now contrast functionalism and supervenience with a causal interpretation of a subjective mental event (see Figure 6c). Here an experimental manipulation causes a nearly instantaneous change in a brain state, which is measured, and this, in turn, later causes a felt subjective mental event, which also is measured. The subjective mental experience, here, has a different etiology than under functionalism (Figure 6a) and supervenience (Figure 6b). In addition, under functionalism, the qualia-like experience is reflected in the relationship between the hypothetical mental event and the subjective mental event (path s in Figure 6a), whereas the qualia-like experience under the causal interpretation is represented in the relationship between the physical brain state and a measured subjective mental state (path u in Figure 6c). But under supervenience, the qualia-like experience is indicated through the relationship between the physical brain state and the hypothetical mental event (path t in Figure 6b). The seemingly similar processes depicted

in Figure 6 constitute subtle differences in the nature of subjective experience and demonstrate that such experience needs to be interpreted within a specific context of theory and method. In addition, note that the correlational relationship (r) between the two outputs of the functional instance (Figure 6a) provides yet a different interpretation of subjective mental events because the observed empirical link between brain state and measured subjective state suggests a connection between a physical state and an expressed qualitative interpretation. Note in Figure 6, we use the letters s, t, and u to refer to distinct origins of qualia-like experiences (a subjective state in Figure 6a, a hypothetical mental state in Figure 6b, and a subjective state in Figure 6c) because the origins of these states and the nature of the concepts linked by s, t, and u are different in the three cases presented in Figure 6. Discussions in the philosophy and neuroscience literatures have sometimes confounded these different meanings of consciousness.

Here, a brief digression is in order concerning the choice of statistical procedures for analyzing data under the models shown in Figure 6 (Antonakis et al., 2010). We thank the Associate Editor for pointing out the key issues here. For the case in Figure 6c, where the mediator and dependent variable are measured, it may be necessary to use instrumental variable estimation. Maximum likelihood or two stage least squares procedures might be used here, for instance. Even in cases where the independent variable is manipulated, and the mediator and dependent variables are measured, bias in inferring causality can occur (e.g., Bullock, Green, & Ha, 2010). Threats to validity still exist, which might require multiple studies to approach each threat in order to increase confidence in findings. For example, support for mediation when mediator and dependent variables are measured in an experiment might be augmented with an additional study where reverse inference from dependent variable to mediator is performed (see Limitations at the end of this article), and still other studies might be done where manipulation of the mediator and its effects on the dependent variable are scrutinized, and where the dependent

variable is switched with the mediator (or even manipulated itself) to see if hypothesized effects are consistent or not with findings. Because hypothetical mental events are present in both Figures 6a and 6b, different statistical methods are needed here, such as ANOVA or multiple regression.

The concept of emergence is another nonreductive interpretation of the nature of mental events. O'Connor and Wong (2015) define emergence as follows: "emergent entities (properties or substances) 'arise' out of more fundamental entities and yet are 'novel' or 'irreducible' with respect to them". Although it has been argued that emergent properties are epiphenomenal and therefore do not enter causal relations and do not even exist (Kim, 1999), other interpretations of emergence allow for physical properties to determine emergent properties, and for emergent properties to influence physical properties and other emergent properties. That said, Nagel (2002) criticizes emergence-based theories as positing a link between two unrelated concepts without making it knowable, and as such explaining nothing, a view many might agree with. It is beyond the scope of this article to consider all the variants, and the limitations and strengths of emergence, so we will limit discussion herein to the related topic of folk psychology.

Folk psychology plays an important role in philosophy of mind and psychology, and one definition that fits our purposes stipulates that folk psychology is "the body of information people have about the mind, and it is often regarded as the basis for our capacity to attribute mental states and to predict and explain actions" (Nichols, 2002, p. 134). Although some philosophers term folk psychology a "bad" theory and recommend that it be "eliminated" (e.g., Churchland, 1981), others call folk psychology a craft and a type of intentional stance useful to people for attributing mental states in others and predicting and influencing the behavior of other people (Dennett, 1987). Whether folk psychology is a theory as discussed extensively in the philosophy literature (see Ravenscroft, 2010) or an instrumental concept akin to a craft, it has a role to play

in organizational science, research, and practice. Folk psychology in everyday behavior often affects people's beliefs, feelings, attitudes, judgment, and behavior (Fletcher, 1995). Moreover, folk psychology often informs theory construction in psychology and applied fields that draw upon psychology (Fletcher, 1995). It is an open question whether folk psychology can be a basis for research into consciousness, from the functionalist perspective or alternatives.

We might envision the use of neuroscience in organizational research from multiple vantage points. Neuroscience may be used to elucidate the nature of important mental phenomena in organizational research. Closely aligned to this purpose, it may be employed to test hypotheses explaining and predicting the behavior of workers and managers. Neuroscience may also be useful in validating and justifying variables and hypotheses such that researchers might draw upon this research to utilize folk psychological variables and theories in basic and applied work, when it is not feasible or prudent to do so by applying neuroscience in such studies.

We believe that models and tests of decision making, goal-directed behaviors, and social action in organizational research can be fruitfully developed based on folk psychology ideas. By grounding folk psychology in philosophy of mind theories and developing frameworks linking the specification of variables and hypotheses with measurement procedures and conceptual criteria in an integrative way, we can gain insight into consciousness and its role in organizational life. Figure 8 presents a framework suggesting the principles advocated here, where for illustrative purposes we sketch a simple theory of action with variables central to folk psychology. Namely, reasons for acting or not are integrated and transformed through a decision maker's desire to act, which, in turn, stimulates an intention to act and then, through intention, action is initiated.

[Figure 8 about here]

Notice in Figure 8 that both mental events and their corresponding physical property concepts are linked to measures and realizers through measurement relationships, which could be done by questionnaire items in the case of subjective reports of mental events and, say, by fMRI in the case of physical brain sates (see the dashed lines, – – – –, in the figure). It should be stressed that measurement relationships, analogous to correspondence rules from the philosophy of science, reflect empirical, theoretical (e.g., theory of the instrumentation), and conceptual criteria describing the relationships between concepts and observations (Bagozzi, 2011b). Moreover, the empirical aspects of the measurement relationships typically contain error (random and systematic), and to the extent that they do, the theoretical/conceptual meaning of the measurements will be contaminated. Thus the measurement links shown in Figure 8 are imperfect in a measurement sense.

Notice further in Figure 8 that the concepts describing the subjective mental events and the concepts describing the physical properties corresponding to them are linked with what we term mind-body relationships (see the solid straight lines in the figure,—). We use the general term, mind-body relationship, simply to convey that the connection in question can be explicated in one of a number of ways. In the case of either pure identity theory or eliminativist perspectives, the mind-body relation is one of identity and the mental and physical collapse into one entity (i.e., each shares all the properties of the other). An identity theorist and a reductive eliminativist might say there are only physical properties (P). An idealist might say there are only ideas (M). But perhaps in the most interesting cases, the mind-body relationship becomes a particular functional specification in a reductive interpretation, or another kind of functional specification in a non-reductive interpretation. An emergent relationship would be another type of non-reductive interpretation. Finally, the mind-body relationship could be interpreted as one of

supervenience. In sum, what we have labelled as the mind-body relationship in Figure 8 for convenience can be interpreted with any one of the major ontological points of view on consciousness. Somewhat analogous to the measurement relationship, the different forms of mind-body relationships do not in conceptual senses perfectly capture the relationships (except of course in the case of identity where the relationship is one of equivalence in the sense of M and P having the same properties). Thus, the nature of the linkages between folk psychology subjective (i.e., qualitative) mental events and the concepts of the physical events has different, imperfect conceptual interpretations, depending on whether one embraces different functional, supervenient, or emergent relationships.

Figure 8 also shows two kinds of causality. The classic and common one occurs between physical concepts (as inferred, for instance, by inferential statistics based on their measurements). These are shown as solid arrows in Figure 8. One physical property, P_1, for example, causes another physical property, P_2 (which would be tested by use of p_1 and p_2). The second kind of causal relation is termed mental causality and is drawn as dashed arrows in Figure 8. We acknowledge that the notion of mental causality is in need of development and is a controversial issue. It should be acknowledged as well that other kinds of causality might occur but are not shown in Figure 8. For example, downward causality might occur from M_1 to P_2, or upward causality from P_2 to M_3. These too are in need of development and are controversial topics.

As a final comment on the general representation of theory of mind and its instantiations shown in Figure 8, we would like to point out the following for perspective: the entire framework in Figure 8 would be necessary for proposing and testing the implications of any valid representation of consciousness. The center and bottom portions of Figure 8 might be used to test physical implications of a theory of mind, where, for instance, Granger causality[2] (e.g., Bressler & Seth, 2011; Kim et al., 2011; Seth, Barrett, & Barnett, 2015) could be applied to fMRI

recordings of activation of P_1-P_4 brain states. The top and center portions of Figure 8 might be used to test subjective state implications of theory of mind, where, for example, multivariate statistics or an experiment could be applied to M_1-M_4 subjective mental concepts and their measurements ($m_1 – m_4$). Of course, to reiterate, the fullest information and test of consciousness requires that the entire system of relationships in Figure 8 be examined in a holistic way (say by an expanded functionalist investigation and the use of structural equation models in an experimental and longitudual study).

NEUROSCIENCE, CONSCIOUSNESS, AND BEYOND

Up to this point, we have placed special emphasis on philosophical accounts of consciousness and have introduced certain concepts, and functional perspectives using these concepts, in order to draw explicit links to foundational issues for the investigation of consciousness by neuroscience. However, empirical research in neuroscience has to date been much narrower in focus and perspective than the topics covered by the philosophical principles we have considered herein. We desire now to more explicitly discuss this orientation by neuroscientists and how it relates to the concepts and principles we have considered[3].

Most neuroscientists appear to be reductionists, and to believe (or at least behave in ways that strongly suggest they implicitly believe) that mental events are fully reducible to physical brain events (e.g., Bickle, 2003). With respect to the question of consciousness more narrowly, there are at least five theories by neuroscientists that fall within the substance monism philosophical tradition we have discussed herein. All five approaches maintain that our brains produce consciousness as a subjective mental state (for short reviews of the theories, see Le Doux, 2015). Consequently, these approaches are consistent with the characterization shown in Figure 6c, where measured brain states lead to or influence subjective experience.

Two of the five theories focus on emotional processes. Panksepp (2012), for example, explains fear behavior as a distillation of subcortical processes in reaction to a threat, where the amygdala influences the hypothalamus which, in turn, influences the periaqueductal gray region of the brain. Fear behavior then results. More elaborate conscious feelings of fear occur also as a function of these subcortical processes in the neocortex. Damasio (2010) also focuses on emotional processes wherein innate emotional behavior and physiological responses become somatic markers to induce basic, conscious feelings in the cortex. For him, information to and from hormones, viscera, and somatosensory receptors in the body interact with both the brainstem (subcortical body-sensing areas) and the thalamus and hypothalamus to undergird what he terms, extended consciousness, which plays out in processes in the insular cortex, prefrontal cortex, and the somatosensory cortex.

The other three theories of consciousness in the brain are somewhat more general and less developed than those proposed by Panksepp (2012) and Damasio (2010), but they set the tone for two, more developed, recent frameworks reviewed below. Sensory processing, especially in the visual cortex, plays a key role in consciousness under all these theories (Le Doux, 2015). For example, first-order theory maintains that consciousness is fully captured in sensory cortical processes (Block, 2007a, b). That is, we process and become aware of a stimulus (phenomenal consciousness), and this leads to a cognitive representation of the state of which we become aware (access consciousness). An elaboration of this point-of-view is access consciousness/attentional amplification theory, wherein prefrontal cortex processes interpret visual cortex processing. Higher-order theory claims consciousness begins with some kind of first-order representation of sensory information (stimulus representation) that is not consciously processed, and then when thought about the stimulus is processed in a higher-order state, leading finally to consciousness (i.e. awareness, Rosenthal, 2012). Last, global workspace theory asserts

that consciousness arises from the to and fro "broadcasting" of information from working memory, on the one hand, and sensory processing, long-term memory, evaluative systems, motor systems, and executive functions (attention, focusing, monitoring) on the other hand (Dehaene & Changeux, 2011).

The theories reviewed above might be better characterized as frameworks, for they have not been specified in detail, operationalized, and tested to demonstrate the production of consciousness. All the general theories specify roles for subcortical processes in consciousness, which play out in one form or another to influence subjective experience. To the best of our knowledge, however, only two programs of neuroscience research to date provide specific consideration of the processes involved at a level of detail, and backed by empirical research, to qualify as theories. Let us consider first the work of Le Doux (2015), for it provides an exemplar for the role of consciousness within the context of a single, albeit complex, emotional state.

Le Doux (2015, p. 11) begins with a distinction between the two related emotions of fear and anxiety that entail overlapping brain mechanisms: "In fear the anticipation concerns if and when a present threat will cause harm, whereas in anxiety the anticipation involves uncertainty about the consequences of a threat that is not present and may not occur". To understand the bases for consciousness here and how it arises, consider Figure 9, which is our summary of most of the major brain regions and their interconnections discussed by Le Doux (2015) in his extensive, ongoing program of research. We discuss first the case of threat. Perception of danger in the form of a threat as an emotional stimulus is processed by the sensory thalamus (see far left center of Figure 9). The thalamus, in turn, initiates parallel processing of a slow sort, beginning in the visual cortex, to take one form of sensory processing as an example, and a faster, nearly automatic, processing by the lateral nucleus of the amygdala (LA). The information in the LA leads quickly to induce behavioral reactions (e.g., freezing, autonomic nervous system (ANS)

activity, hormonal production), through its interactions with the central nucleus (CeA) and the basal nucleus (BA) of the amygdala. Processing by the LA also leads to actions (i.e., such motor responses as escape or avoidance) through the pathway of BA to the bed nucleus of the stria terminalis (BNST), to the nucleus accumbens (NAcc), to the ventral pallidum (VP). These automatic-like paths of threat processing occur sub-cortically, and although we later can become aware of our freezing, ANS, hormal, and motor responses, they are not conscious processes. This fast processing route to perceptions of danger, of course, functions to enhance survival in non-deliberative ways when needed.

[Figure 9 about here]

The slow route to processing threat occurs primarily in higher cortical regions as shown in Figure 9, and is useful when more detailed, fine-grained thinking functions are needed to deal with danger. The visual cortex begins this activity, where one path eventually gets to the LA through the tertiary visual cortex, and so also can initiate behavioral reactions and motor responses indirectly. Le Doux (2015) stresses that the amygdala has indirect effects on cortical processing via the CeA and arousal systems, and direct effects on cortical processing via the BA. By arousal systems, Le Doux (2015) means neurons that produce such neuromodulators as norepinephrine, serotonin, dopamine, acetylcholine, orexins, and others. He claims that the feedback loops (see Figure 9) lead to amplification of processing and ultimately promote the emergence of consciousness. Feelings for Le Doux (2015, p. 228), in this case fear, get "their characteristic quality (the way they feel)…by the combination of non-emotional ingredients". That is, feelings are peculiar mixtures of sensory processing, brain arousal, survival circuit activity (ANS, hormones), feedback from behavioral and physiological reactions, memory, and executive functions (attention, monitoring, labeling, attributing). This occurs in working memory in the lateral prefrontal cortex and draws upon long-term memory in the hippocampus as well.

Anxiety follows similar dynamics, as well as certain distinct activities: "Like fear, anxiety can be initiated from the outside…[but] can also be triggered by stimuli that are only weakly associated with danger…Novel situations in which one does not know exactly what to expect are also triggers of anxiety…[and] anxiety can develop when particular memories or thought…lead to worry" (Le Doux, 2015, p. 231). Our qualitative sense of what it feels like to experience fear or anxiety are produced, Le Doux (2015) asserts, in working memory to give us such fine-grained distinctions as anxiety, apprehensiveness, concern, creepiness, dread, fear, fright, horror, nervousness, panic, terror, feeling troubled, and worry, depending on the amalgam of cortical and sub-cortical information integrated to produce them.

Le Doux's (2015) research shows how a single mental event, fear/anxiety, is produced by cortical and subcortical process to shape subjective awareness of these mental representations. It is one of the most developed treatments in the literature and only recently appeared in the form described above. It falls squarely in a functionalist framework and has been studied this way in terms of conditioning and extinction experiments. Fear/anxiety in organizational research is an important component of stress in and of itself (e.g., in role conflict) or as part of larger theories for coping with stress (e.g., attachment theory). Le Doux's (2015) research could be paradigmatic for studying conscious processes of fear/anxiety in organizational research, or for other emotions or constructs. However, although Le Doux (2015) claims that the cortical and subcortical processes he has studied produce consciousness (as shown in Figure 6c), this has not been tested directly. Most neuroscience research has been of the sort more closely corresponding to Figure 6b or Figure 6a, where actual subjective mental events are not included or measured. Multilevel research is needed to more directly explore the role of consciousness in neuroscience (see Figure 5). Here, however, the approaches taken to date have shown that (a) hypothesized mental events play an intermediary role explaining both physical brain states and subjective brain states and (b)

physical brain states and subjective brain states, dependent on the same intermediary mental state, are highly correlated. Whether the correlations are indicative of brain states producing consciousness (as in Figure 6c), or consciousness causing brain states (where both are not causally related but are associated because of common causes, or express some other relationship) remains to be determined.

A final paradigmatic case for study of consciousness in neuroscience goes beyond the investigation of a single mental event, such as emotion, to encompass two fundamental processes underpinning much of organizational life. The psycho-social reward system in the brain depends on the coordination of several brain regions: "reward processing involves multiple components, including one that is motivational (*wanting*) and another that is hedonic (*liking*) which rely on separate neural networks that can be dissociated under particular circumstance" (Pool et al., 2016, p. 126). Common sense and many theories in psychology and organizational studies have long assumed that the things we desire are things we like, and the things we like are things we desire. But recent research in neuroscience has shown that wanting and liking follow different etiologies. So for example, classic understandings of work commitment and job satisfaction may have to be reconceived to discover how, or even whether, these two fundamental processes relate to each other in particular contexts.

Motivation has been specified to play out in the wanting neural network, in which it occurs in implicit and explicit forms (e.g., Berridge & Robinson, 2003; Anselme & Robinson, 2015). The former is basically a Pavlovian-based system of learning where unconditioned stimuli are paired with conditioned stimuli to comprise an automatic, unconscious motivation. Employees, for instance, learn to value rewards inherent in performing certain work tasks that are, in turn, linked to sought-for outcomes. Wanting in this sense adds "a visceral oomph to mental desires,…corresponds best to decision utility,…[and] could be a mechanism for decision

that is distinguishable from both experienced utility (hedonic impact or "liking" the outcome) and forecast or predicted utility (expecting in advance to like an eventual outcome)" (Berridge, 2009, pp. 378-379). Wanting also underlies a conscious, goal-directed system. It is this system that unconscious desires "color…with motivational power, to make them compelling spurs to action" (Berridge, 2009, p. 379).

The other component of the psycho-social reward system captures the hedonic experiences (liking) of the rewards employees achieve. Unlike the wanting system that depends heavily on dopamine effects in the brain and involves the ventral tegmental area, NAcc, anaygdala, insular cortex, VP, and orbitofrontal cortex, the liking system is more localized in a small number of brain regions, termed hedonic hotspots, where opioids and endocannabinoids function (e.g., Salamone and Correa, 2012). The liking system also overlaps with the wanting system in terms of involvement of the NAcc, VP, and orbitofrontal cortex.

The wanting and liking systems are largely nonconscious. But researchers maintain that conscious pleasure is a function of both systems (e.g., Kringelback & Berridge, 2012). We might speculate that the nonconscious processes lead to awareness of one's desires and likes and dislikes along the lines suggested by Le Doux (2015) for fear/anxiety. However, production of consciousness as a function of wanting and liking is in need of verification and has not been studied systematically. The possibility of consciousness influencing wanting, liking, and other neurosystems should also be considered.

The fear/anxiety and the wanting/liking research are early examplars, we suggest, for study of how conscious phenomena function in organizational contexts. Lieberman (2010) has summarized neuroscience research that can serve as foreshadowings of promising areas for substrates of important psychological and social processes, but none of this research has yet to explain conscious processes as outcomes of physical brain states to the degree that Le Doux

(2015) and the other researchers mentioned above have done. The areas with neuroscience foundations that Lieberman (2010) reviews are social perception, social inference, self-processes, social interaction, attitudes and attitude change, and stereotyping and intergroup processes, wherein each of these categories has numerous subcategories such as empathy, mentalizing, mirroring, morality, agency, self-control, self-concept, trust and cooperation, altruism, and prejudice. There are obviously many opportunities for research in organizational behavior that neuroscience concepts and methods can inform. There is need for conceptualization and research into consciousness that construes it as either a product of brain processes or constituted by such processes in the sense discussed above by Nagel (2002, 2012), or possibly a determinant of brain processes. We hope that the philosophical foundation and methodological framework provided herein can help organizational researchers launch new research endeavors with the sound conceptual tools needed for success in this regard.

LIMITATIONS AND FUTURE DIRECTIONS

The literature on consciousness and neuroscience is vast, and a number of important issues could not be addressed herein. One of these is the notion of reverse inference (see Lee, Senior, and Butler, 2012), where the goal is "to infer the likelihood of a particular mental process M from a pattern of brain activity A, which can be framed as a conditional probability P(M/A)", and Bayes' rule applied (Poldrack, 2011, p. 693). Poldrack (2011) reviews recent advances in reverse inference, including a rebuttal concerning criticisms of the approach. Whereas discussion of the issues here are beyond the scope of our article, it should be noted that the role and interpretation of causality are different between the functional approaches considered herein, on the one hand, where consciousness gets is meaning from the hypothesized role it plays between a causal manipulation and observed behavior, and reverse inference, on the other hand, where base rates are needed and the likelihood of a mental event is to be inferred from brain activity. The

statistical assumptions between Bayesian and Fisherian/frequentist approaches also entail conceptual and practical differences that need to be considered in order to comprehend the issues fully and make informed research choices, and as such should be topics for future research.

A second important issue that was not possible to consider herein is the relationship of the critical perspective in management and social science to neuroscience research. The critical perspective in management studies, which entails a massive body of work and follows multiple points of view, might be characterized as "a combination of Frankfurt School/radical humanism and critical postmodernism" and has "shown how managerial values embedded in language systems, social practices, and decision routines have lessened the quality of organizational decisions and reduced the capacity to meet important human needs" (Alvesson & Deetz, 2006, p. 255). We could find no research explicitly relating the critical management perspective to neuroscience, although Lindebaum's (e.g. 2016; Lindebaum and Zundel, 2013) criticism of organizational neuroscience does appear to share some commonalities in perspective. Even so, we speculate that there are at least two areas for future enquiry that go beyond the simple critique of reductionist neuroscience.

One area is largely normative and addresses social, cultural, and political issues with regard to critical practice (Choudhury, Nagel, & Slaby, 2009). Here implications of neuroscience in organizations can be scrutinized to analyse contextual and historical factors with current neuroscience research so as to consider what has been termed, "the central goal of critical theory" in organizational studies: namely, "to create societies and work places which are free from domination, where all members have an equal opportunity to contribute to the production of systems that meet human needs and lead to the progressive development of all" (Alvesson & Deetz, 2006, p. 259). Such issues as public trust, neuroethics, moral decision making, the role of funding agencies, and the assessment of scientific quality are a few concerns for study here (e.g.,

Schleim, 2014). Lindebaum (2016) touches on a number of these issues, but existing critical neuroscience work has gone considerably further, and there seems ample opportunity to draw from, and contribute to, such discussions by critical management scholars.

The second area that the critical perspective might inform neuroscience research in organizations is with regard to the conceptualization of consciousness and how it is investigated. Specifically, what is the proper representation of consciousness and its role in management and organization studies? Here a critical realist orientation might yield balanced guidelines for enquiry (e.g., Healey & Hodgkinson, 2014). Rather than subscribing to a narrow reductive physicalism for thinking about consciousness, ideas roughly falling under the rubric of the social construction of reality might be better accommodated by use of such orientations as emergentism, non-reductive functionalism, dualism, downward causation, and folk psychology, so as to provide ideas for translating concepts and prescriptions from the critical turn in organization research into neuroscience enquiry in organizations. The role of embodied cognition and the use of metaphors in interpretation, expression, and construction of subjective meaning deserve more consideration as well if we are to get a sense of what consciousness is.

To conclude, the meaning of consciousness is embedded in seemingly opposite or irreconcilable and entrenched points of view. A key question underlying all discussions of the meaning of consciousness is whether anything beyond the physical exists, such as transcendent, supernatural or natural or subjective non-physical entities. Researchers on both sides of these issues tend to push inquiry to the extremities with little room for dialogue or reconciliation. We suggest that a third path is needed to open-up discussion and consider whether consciousness in some sense is unique, whether it sometimes emerges out of physical processes, and even whether it sometimes influences physical processes. Our aim in this article has been to consider multiple vantage points so as to avoid strident and premature commitments to one perspective or another.

But it will take continual struggle and tolerance, we suspect, to arrive at workable frameworks for consciousness. Here, we submit, the Principles of Charity and Relevant Evidence (see Glossary) from philosophy could be followed by all concerned.

Footnotes

1. Although he is never explicit, Nagel (2002, 2012) himself occasionally implies that he considers his position to be a form of "neutral monism" (e.g. 2012, pp. 4-5, see also Sorrell, 2005). Even so, it seems that a form of property dualism fits Nagel's key theses more accurately, being that traditional neutral monism takes a position where the nature of reality is *neither* mental nor physical (Stubenberg, 2016). Conversely, Nagel's position favors the view that the nature of reality is *both*. Certainly, some forms of neutral monism take the same position either implicitly or explicitly, but in such cases this would seem to place them more in line with a property / naturalistic dualist position (see Stubenberg, 2016).

2. Granger causality is hard to explain briefly, and interested readers are referred to the cited works in relation to neuroscience, and also to Granger (2004) and Gujarati & Porter (2008) in a more general sense. In essence, Granger causality refers to the ability of one time series to predict another, rather than the correlation approach of ordinary regression. Intuitively-speaking, if x and y are both time-evolving variables, one says that x Granger-causes y if the predictions of y based on its own past values and the past values of x are better than predictions of y based its past values alone. Given the philosophical complexity of defining what causality actually is, Granger causality is generally considered to find only predictive causality.

3. By way of further explication, the present discussion was inspired in part by early presentations of draft versions of this article. On more than one occasion, neuroscientists raised their hands to announce "this is not neuroscience".

References

Alvesson, M., & Deetz, S.A. (2006). Critical theory and postmodernism approaches to organization studies. In S.R. Clegg, C. Hardy, T.B. Lawrence, & W.R. Nord (Eds). *The Sage Handbook of Organization Studies*, (pp. 255-283). Thousand Oaks, CA: Sage.

Anselme, P., & Robinson, M.J. (2015). "Wanting", "liking", and their relation to consciousness. *Journal of Experimental Psychology Animal Learning and Cognition*, 42, 123-140.

Ashkanasy, N. M., Becker, W. J., and Waldman, D. A. (2014). Neuroscience and organizational behavior: Avoiding both neuro-euphoria and neuro-phobia. *Journal of Organizational Behavior*, 35(7):909-919.

Antonakis, J., Bendohan, S., Jacquart, P., & Lalive, R. (2010). On making casual claims: A review and recommendations. *The Leadership Quarterly*, 21, 1086-1120.

Bagozzi, R.P. (2011a). Alternative perspectives in philosophy of mind and their relationship to structural equation models in psychology. *Psychological Inquiry,* 22, 88-99.

Bagozzi, R.P. (2011b). Measurement and meaning in information systems and organizational research: Methodological and philosophical foundations. *MIS Quarterly*, 35, 261-292.

Bagozzi, R.P., Verbeke, J.M.I., van den Berg, W., and Rietdijk, Dietvorst, R.C., and Worm, L. (2012). Genetic and neurological foundations of customer orientation: Field and experimental evidence. *Journal of the Academy of Marketing Science*, 40, 639-658.

Bagozzi, R.P., Verbeke, J.M.I., Dietvorst, R.C., Belschak, F.D., van den Berg, W., and Rietdijk, W.J. R. (2013). Theory of mind and empathic explanations of Machiavellianism: A neuroscience perspective. *Journal of Management*, 39, 1760-1798.

Baron-Cohen, S. (2011). *The science of evil: ON empathy and the originas of cruelty*. New ork: Basis Books.

Barrett, L.F. (2011). Bridging token identity theory and supervenience theory through psychological construction. *Psychological Inquiry*, 22,115-127.

Becker, W.J., Cropanzano, R., and Sanfey, A.G. (2011). Organizational neuroscience: Taking organizational theory inside the neural black box. *Journal of Management*, 37, 933-961.

Berridge, K.C. (2009). Wanting and liking: Observations from the neuroscience and psychology laboratory. *Inquiry*, 52, 378-398.

Bickle, J. (2003). *Philosophy and neuroscience: A ruthlessly reductive account*. Dordrecht: Kluever.

Blackburn, S. (1994). *The Oxford dictionary of philosophy*. Oxford: Oxford University Press.

Block, N. (2007a). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, 30, 481-548.

Block, N. (2007b). *Consciousness, function, and representation*. MIT Press.

Braddon-Mitchell, S. and Jackson, F. (2007). *Philosophy of mind and cognition*, 2nd ed., Malden, MA: Blackwell.

Bressler, S.L., & Seth, A.K. (2011). Wiener-Granger casuality: A well established methodology. *Neuroimage*, 58, 329-329.

Bullock, J.G., Green, D.P., & Ha, S.E. (2010). Yes, but what's the mechanism? (Don't expect an easy answer). *Journal of Personality and Social Psychology*, 98, 550-558.

Burge, T. (2007). *Foundations of mind: Philosophical essays*. Vol 2. Oxford: Oxford University Press.

Cacioppo, J.T., Bernston, G.G. (1992). Social psychological contributions to the decade of the brain. Doctrine of multilevel analysis. *American Psychologist*, 47, 1019-1028.

Cacioppo, J.T., and Decety, J. (2009). What are the brain mechanisms on which psychological processes are based? *Perspectives on Psychological Science*, 4, 10-18.

Cacioppo, J. T. and Decety, J. (2011). Social neuroscience: challenges and opportunities in the study of complex behavior. *Annals of the New York Academy of Sciences*, 1224(1):162-173.

Cacioppo, J.T., Bermiton, G.G., Sherida, J.F., & McClintock, M.K. (2000). Multilevel integrative analysis of human behaviour: Social neuroscience and the complementing nature of social and biological approaches. *Psychological Bulletin*, 126, 829-843.

Cartwright, N. (1999). *The Dappled World*. Cambridge: *Cambridge University Press.*

Chalmers, D.J. (2010). The character of consciousness. Oxford: *Oxford University Press.*

Choudhury, S. & Slaby, J. eds. (2011). *Critical neuroscience: A handbook of the social and cultural contexts of neuroscience*. New York; Wiley-Blackwell.

Choudhury, S., Nagel, S.K., & Slaby, J. (2009). Critical neuroscience: Linking neuroscience and society through critical practice. *BioSocieties*, 4, 61-77.

Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *Journal of Philosophy*, 78(2):67-90.

Crick, F.J. (1994). *The astonishing hypothesis: The scientific search for the soul*. New York: Scribner's.

Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, 70, 200-227.

Damasio, A.R. (1994). *Descartes' error: Emotion, reason and the human brain*. New York: Avon books.

Damasio, A. (2010). *Self comes to mind*. New York: Vintage.

Dennett, D. (1987). True believers: The intentional strategy and why it works. In D. Dennett, *The Intentional Stance*. Cambridge, MA: MIT Press, pp. 14-35.

Dennett, D. (1991). *Consciousness explained*. Boston: Little, Brown.

Dennett, D. (2003). *Freedom evolves*. London: Allen Lane Books.

Dennett, D. (2006). *Sweet dreams: Philosophical obstacles to a science of consciousness*. Cambridge, MA: MIT Press.

Dennett, D. (2013). *Intuition pumps and other tools for thinking*. London: Allen Lane.

Dietvorst, R.C., Verbeke, W.J.M.I., Bagozzi, R.P., Yoon, C., Smits, M., and van der Lugt, A. (2009). A sales force-specific theory-of-mind scale: Tests of its validity by classical methods and functional magnetic resonance imaging. *Journal of Marketing Research*, 46, 653-668.

Feyerabend, P. K. (1963). Comment: Mental events and the brain. *The Journal of Philosophy*, 60(11):295-296.

Fletcher, G.J.O. (1995). Two uses of folk psychology: Implications for psychological science. *Philosophical Psychology*, 8,221-238.

Franks, D.D. (2013). Emergence and reductionism in sociology and neuroscience. In D.D. Franks and J.H. Turner (Eds.), *Handbook of Neurosociology*. New York: Springer, 107-116.

Fodor, J (1990). *A Theory of Content and Other Essays*, Cambridge, Mass, Bradford Book/MIT Press.

Gazzaniga, M.S. (2011). Who's in charge?: Free will and the science of the brain. New York: HarperCollins.

Granger, C.W. (2004). Time series analysis, cointegration, and applications. *American Economic Review*, 94 (3): 421-425.

Gray, J. A. (1995). The contents of consciousness: A neuropsychological conjecture. *Behavioral and Brain Sciences*, 18:659-676

Gujarati, D. & Porter, D. (2008). *Basic Econometrics (5ed)*. McGraw-Hill, New York.

Gutting, G. (2015). *What Philosophy Can Do*. W. W. Norton and Co., New York.

Guyer, P. & Horstmann, R.-P. (2015). Idealism, in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition), URL = <http://plato.stanford.edu/archives/fall2015/entries/idealism/>.

Hawking, S. and Mlodinow, L. (2010). *The Grand Design*. Bantam, London, UK

Healey, M. P. and Hodgkinson, G. P. (2014). Rethinking the philosophical and theoretical foundations of organizational neuroscience: A critical realist alternative. Human Relations, 67(7):765-792.

Kievit, R.A., Romeign, J-W, Waldorp, L.J., Wicherts, J.M., Scholte, H.S., and Borsboom, D. (2011a). Mind the gap: A psychometric approach to the Reduction Problem. *Psychological Inquiry,* 22, 67-87.

Kievit, R.A., Romeign, J-W, Waldorp, L.J., Wicherts, J.M., Scholte, H.S., and Borsboom, D. (2011b). Modelling mind and matter: Reductionism and psychological measurement in cognitive neuroscience. *Psychological Inquiry*, 22, 139-157.

Kim, J. (1993). *Supervenience and Mind: Selected Philosophical Essays*. Cambridge University Press, Cambridge, UK

Kim, J. (1999). Making sense of emergence. *Philosophical Studies*, 95, 3-36.

Kim, J. (2000). *Mind in a physical world: An essay on the mind-body problem and mental causation*. Cambridge, MA: MIT Press.

Kim, J. (2005). *Physicalism, or something near enough*. Princeton, NJ: Princeton University Press.

Kim, S., Putrino, D., Ghosh, S. & Brown, E.N. (2011) A Granger Causality Measure for Point Process Models of Ensemble Neural Spiking Activity. *PLOS Computational Biology*, http://dx.doi.org/10.1371/journal.pcbi.1001110

Kripke, S. (1980). *Naming and necessity*. Oxford, UK: Blackwell.

Kosslyn, S. M. (1999). If neuroimaging is the answer, what is the question? *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 354(1387):1283-1294

Lamm, C., Batson, C. D., and Decety, J. (2007). The neural substrate of human empathy: Effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience*, 19(1):42-58

Le Doux, J. (2015). *Anxious*. New York: Viking.

Lee, N., Senior, C., and Butler, M.J.R. (2012). The domain of organizational cognitive neuroscience: Theoretical and empirical challenges. *Journal of Management*, 38 (4), 921-931.

Levin, J. (2013), Functionalism, in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), URL = <http://plato.stanford.edu/archives/fall2013/entries/functionalism/>

Lieberman, M.D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, 58, 259-289.

Lieberman, M. D. (2010). Social cognitive neuroscience. In *Handbook of Social Psychology*. John Wiley & Sons, Inc

Lindebaum, D. (2016). Critical essay: Building new management theories on sound data? the case of neuroscience. *Human Relations*, 66(6):1-14

Lindebaum D. and Zundel M. (2013). Not quite a revolution: Scrutinizing organizational neuroscience in leadership studies. *Human Relations,* 66 (6), 857-877.

Marraffa, M., Theory of mind, in *The Internet Encyclopedia of Philosophy*, ISSN 2161-0002, http://www.iep.utm.edu/theomind, accessed on 14 Feb, 2016.

McLaughlin, B. & Bennett, K, (2014), Supervenience, in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), URL = <http://plato.stanford.edu/archives/spr2014/entries/supervenience/>.

Nagel, T. (2002). *Concealment and Exposure: And Other Essays*. Oxford University Press, Oxford, UK.

Nagel, T. (2012). Mind and Cosmos: Why the materialist neo-Darwinian conception of nature is almost certainly false. New York: Oxford University Press USA

Nichols, S. (2002). Folk psychology. In *Encyclopedia of Cognitive Science*, London: Nature Publishing Group, pp. 134-140.

O'Connor, T. and Wong, H.Y. (2015). Emergent Properties, in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), URL = <http://plato.stanford.edu/archives/sum2015/entries/properties-emergent/>.

Panksepp, J. (2012). *The archaeology of mind: Neuroevolutionary origin of human emotions*. New York: Norton.

Pinker, S. (2015), How the mind works. London: Penguin Random House UK.

Poldrack, R.S. (2011). Inferring mental states from neuroimaging data: From reverse inference to large-scale decoding. *Neuron*, 72, 692-697.

Pool, E., Semmwald, V., Delplanque, S., Brosch, T., & Sander, D. (2016). Measuring wanting and liking from animals to humans: A systematic review. *Neuroscience & Biobehavioral Reviews*, 63, 124-142.

Pozharliev, R., Verbeke, W. J. M. I., Van Strien, J. W., and Bagozzi, R. P. (2015). Merely being with you increases my attention to luxury products: Using EEG to understand consumers' emotional experience with luxury branded products. *Journal of Marketing Research*, 52(4):546-558

Ramsey, W. (2013), Eliminative Materialism, in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2013 Edition), URL = <http://plato.stanford.edu/archives/sum2013/entries/materialism-eliminative/>.

Ravenscroft, I. (2016), Folk Psychology as a Theory, in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2016 Edition), forthcoming URL = <http://plato.stanford.edu/archives/fall2016/entries/folkpsych-theory/>.

Robinson, H. (2016). Dualism, in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), URL = <http://plato.stanford.edu/archives/spr2016/entries/dualism/>.

Rosenthal, D. (2012). Higher-order awareness, misrepresentation and function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 1424-1438

Salvador, R. and Folger, R. G. (2009). Business ethics and the brain. *Business Ethics Quarterly*, 19:1-31

Schaffer, J. (2016). Monism, in E.N. Zalta (ed.),, *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), URL = <http://plato.stanford.edu/archives/spr2016/entries/monism/>.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., and Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, 36:393-414.

Schleim, S. (2014). Critical neuroscience--or critical science? A perspective on the perceived normative significance of neuroscience. *Frontiers in Human Neuroscience*, 8, article 336, 1-6.

Senior, C., Lee, N., and Butler, M.J.R. (2011). Organizational cognitive neuroscience. *Organizational Science*, 22, 804-815.

Seth, A.K., Barrett, A.B.,& Barnett, L. (2015). Granger causality analysis in neuroscience and neuroimaging. *Journal of Neuroscience*, 35(8): 3293-3297.

Silbersweig, D. (2015). A Harvard medical school professor makes the case for the liberal arts and philosophy. *Washington Post*, December 24, 2015.

Smart, J. J. C. (2014). The Mind/Brain Identity Theory, in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), URL = <http://plato.stanford.edu/archives/win2014/entries/mind-identity/>.

Stack, G. J. (2000). Materialism, in *The Concise Routledge Encyclopedia of Philosophy,* Routledge, London, UK.

Stoljar, D. (2016). Physicalism, in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Spring 2016 Edition)*, URL = http://plato.stanford.edu/archives/spr2016/entries/physicalism/

Stubenberg, L. (2016). Neutral Monism, in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), forthcoming URL = <http://plato.stanford.edu/archives/win2016/entries/neutral-monism/>.

Tam, L., Bagozzi, R.P., & Spanjol, J. (2010). When planning is not enough: The self-regulatory effect of implementation intentions on changing snacking habits. *Health Psychology*, 29, 294-292.

Tellis, R. (2011). *Aping mankind: Neuromania, Darwinitis, and the misrepresentation of humanity*. New York: Routledge.

van Riel, R. and Van Gulick, R. (2016). Scientific Reduction, in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), URL = http://plato.stanford.edu/archives/spr2016/entries/scientific-reduction/

Verbeke, W., Bagozzi, R.P., and van den Berg, W.E. (2014). The role of attachment styles in regulating the effects of dopamine on the behavior of salespersons. *Frontiers in Human Neuroscience*, 04 February, 18, 1-13.

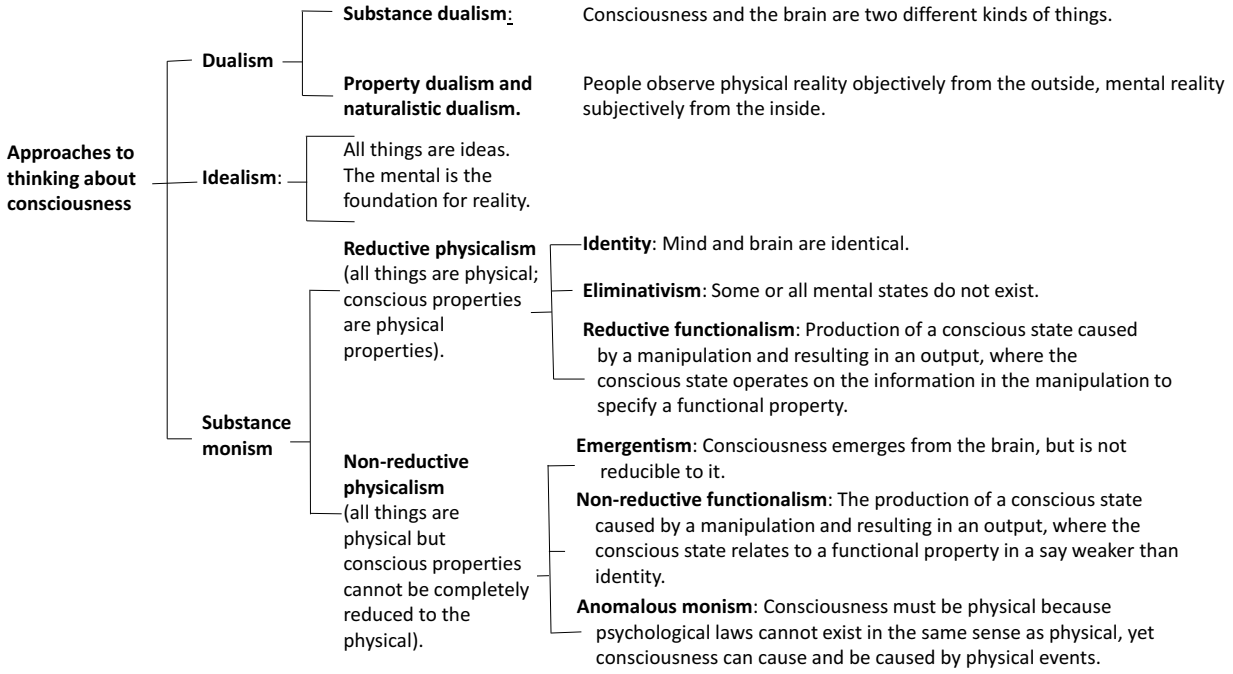
Waldman, D. A., Balthazard, P. A., and Peterson, S. J. (2011). Leadership and neuroscience: Can we revolutionize the way that inspirational leaders are identified and developed? *The Academy of Management Perspectives*, 25(1):60-74

Wegner, D. M. (2002). *The Illusion of Conscious Will*. MIT Press, Cambridge, MA

Wright, R. (2000). *Nonzero: The Logic of Human Destiny*. Pantheon, New York,

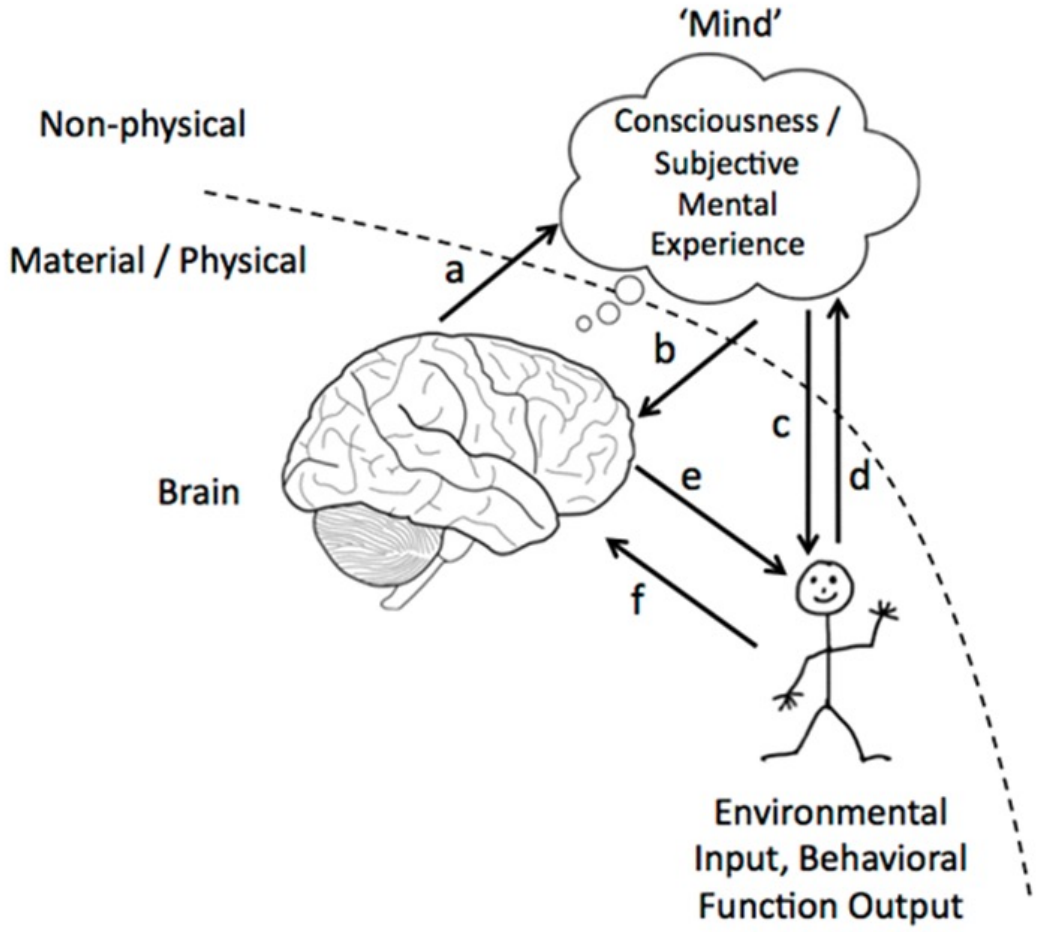
Yalowitz, S. (2014), Anomalous Monism, in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), URL = <http://plato.stanford.edu/archives/win2014/entries/anomalous-monism/>

FIGURE 1: PHILOSOPHICAL APPROACHES TO THINKING ABOUT CONSCIOUSNESS

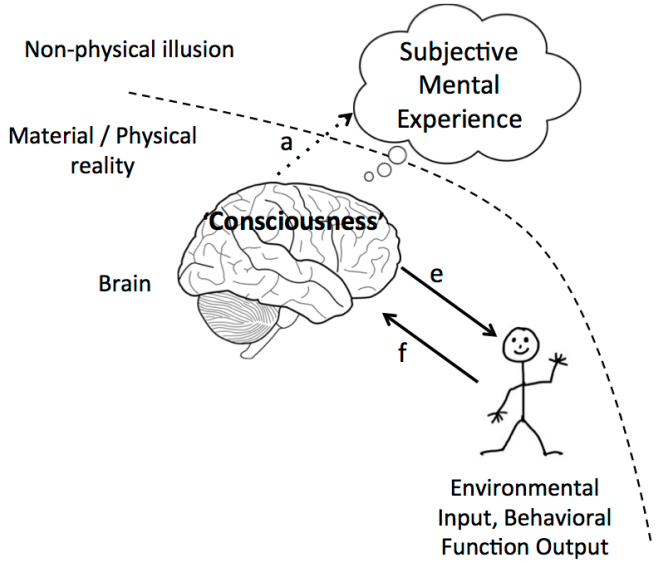


Note: For more extensive definitions and historical commentary, see appendix.

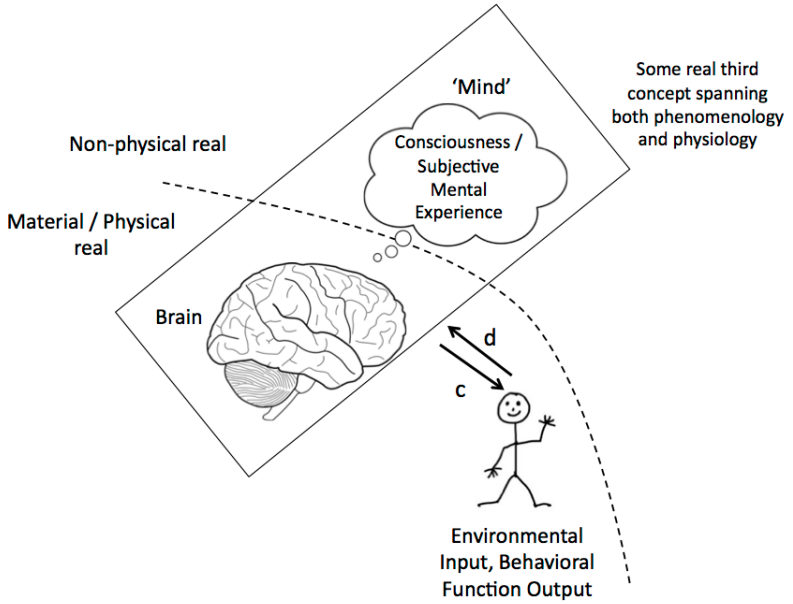
FIGURE 2: A NAÏVE MODEL OF THE BRAIN-MIND-BEHAVIOR LINK



FIGURES 3: COMPARISON OF PHYSICALIST AND NONPHYSICALIST PRINCIPLES

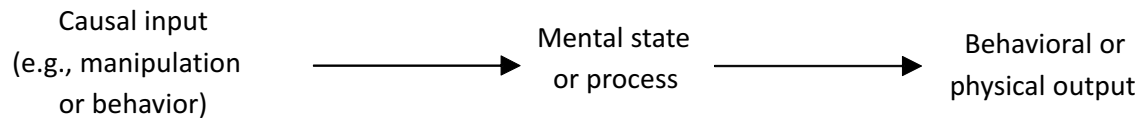


3A: PHYSICALIST

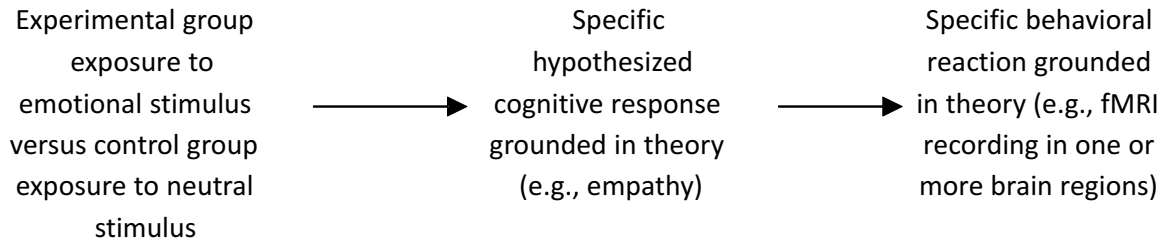


3B: NONPHYSICALIST

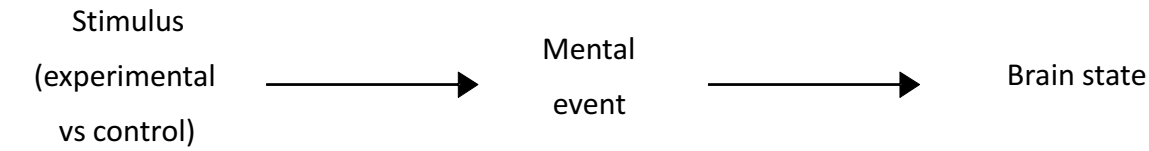
FIGURE 4. FUNCTIONALISM



4a. General form of functionalism



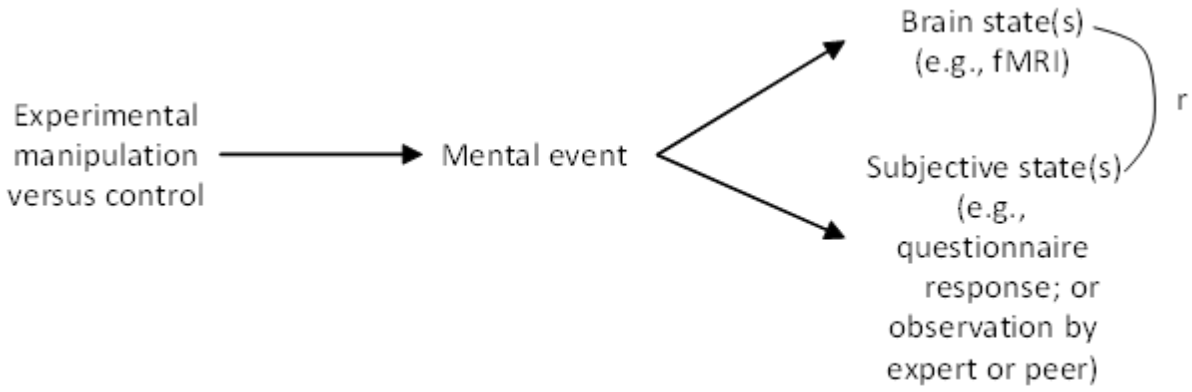
4b. Example of functionalism



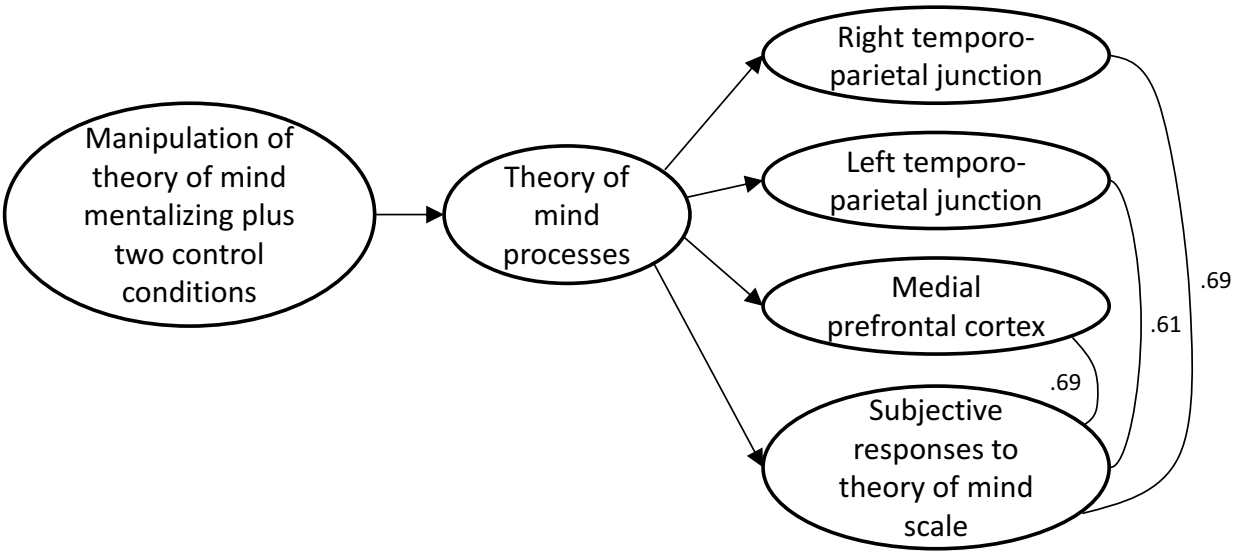
General form of

4c. Functionalism in Neuroscience

FIGURE 5: MULTILEVEL RESEARCH WITHIN THE CONTEXT OF FUNCTIONAL EXPLANATIONS

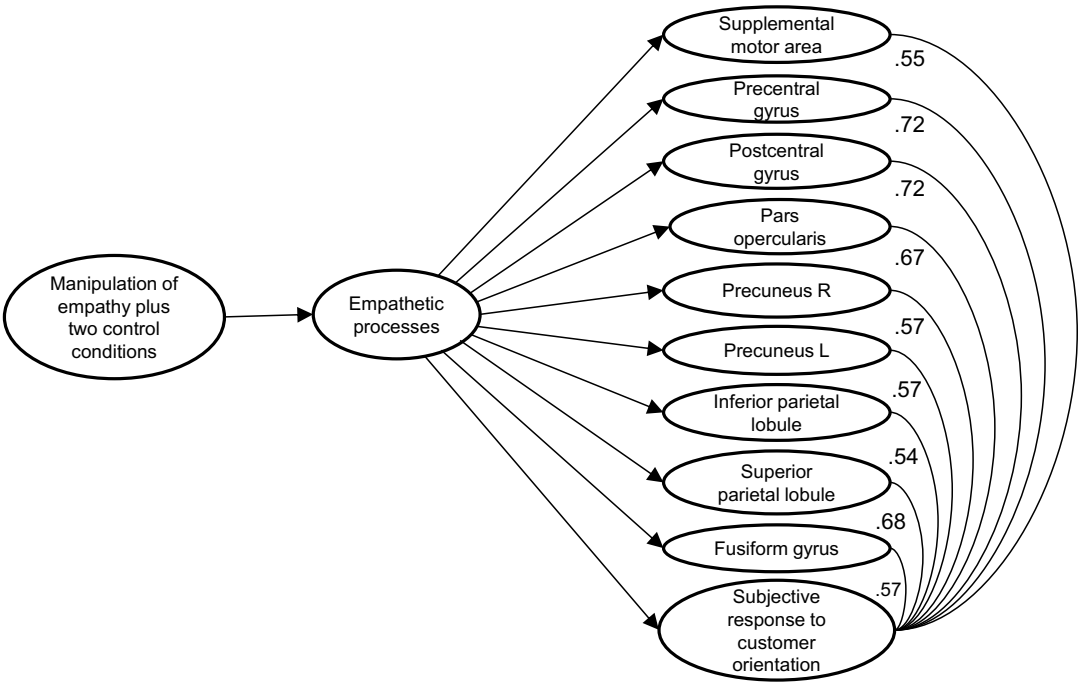


5a. Multilevel research with two criterion levels

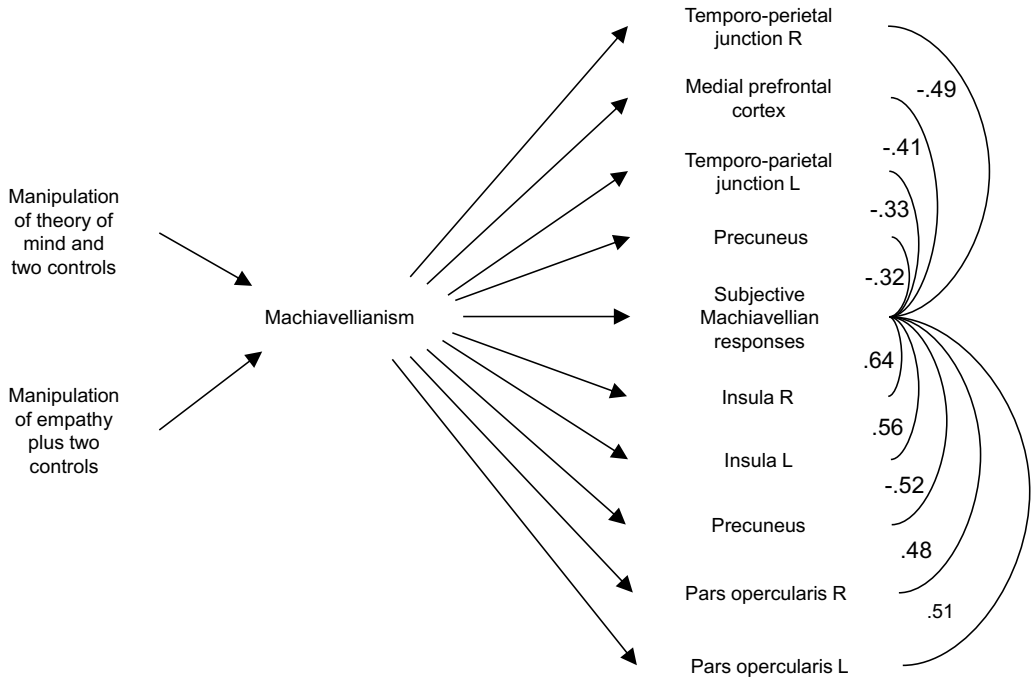


5b. Theory of mind study with brain states (fMRI) and subjective reports on a psychological scale

FIGURE 5: CONTINUED

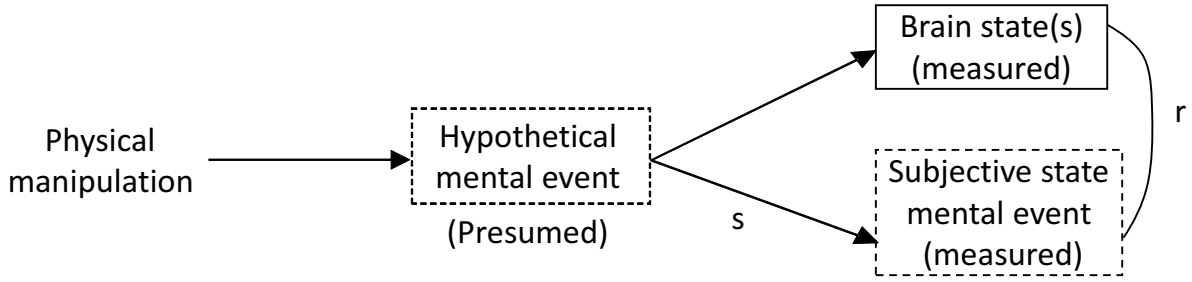


5c. Empathy and mirror neurons with brain states and subjective reports of implications of empathy (i.e., customer orientation)

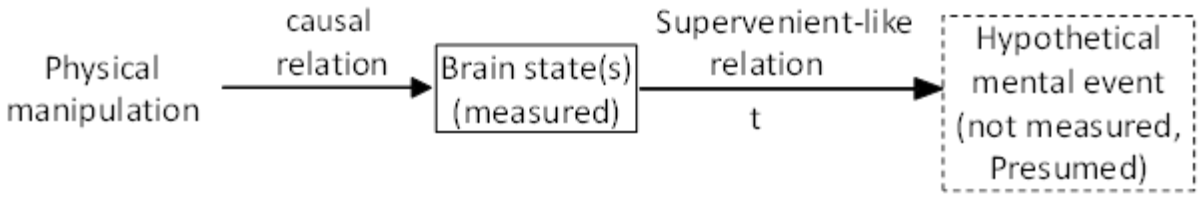


5d. Machiavellianism

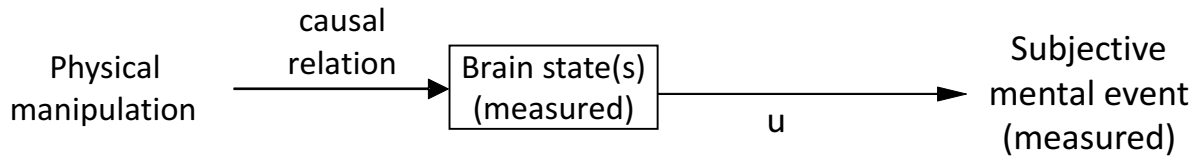
FIGURE 6. COMPARISONS AND CONTRASTS BETWEEN FUNCTIONAL, SUPERVENIENT, AND CAUSAL RELATIONSHIPS FOR HYPOTHETICAL VERSUS SUBJECTIVE MENTAL EVENTS



6a. Functional interpretation of hypothetical mental event



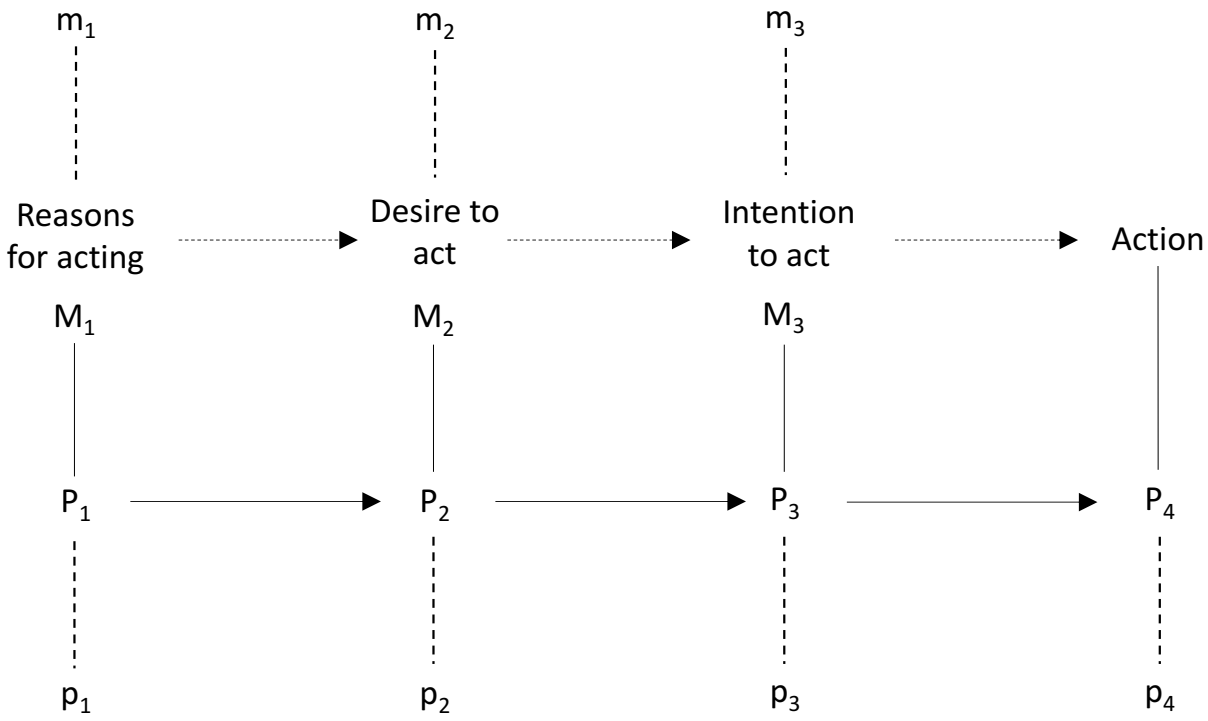
6b. Supervenient interpretation of hypothetical mental event



6c. Causal interpretation of measured mental event

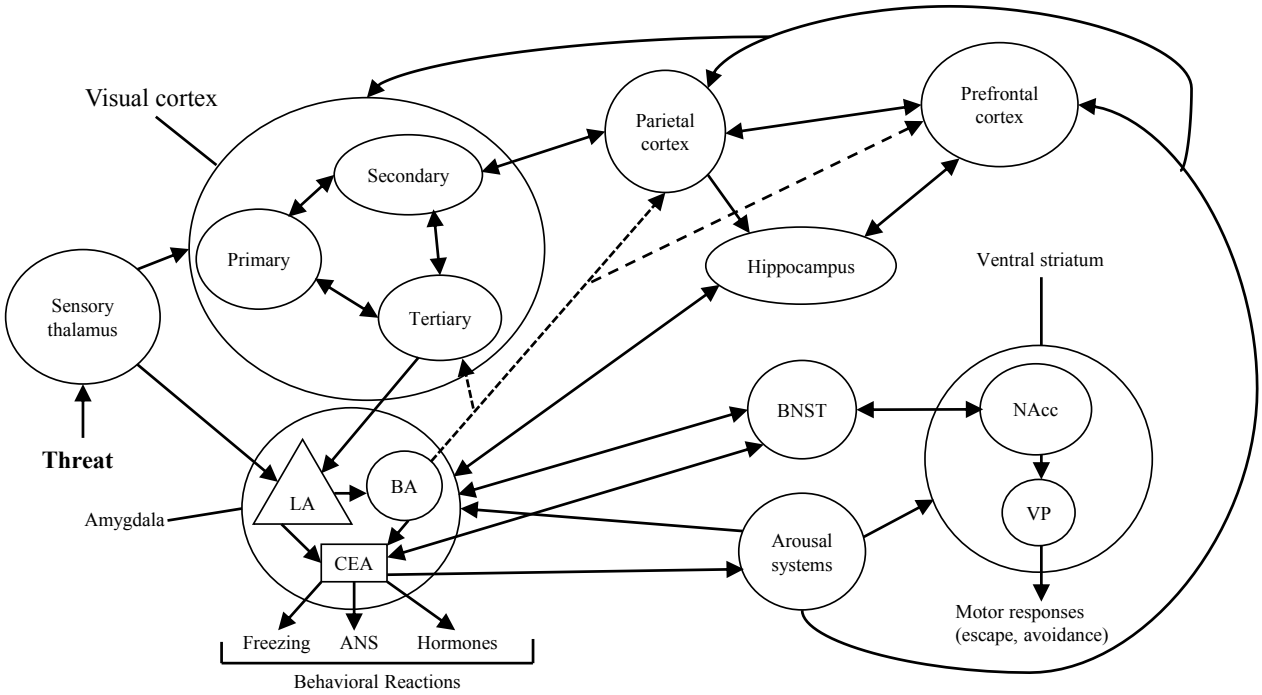
FIGURE 7 SUPERVENIENCE EXAMPLE

FIGURE 8. GENERAL REPRESENTATION OF PHILOSOPHY OF MIND APPLIED TO A FOLK PSYCHOLOGY EXPLANATION OF ACTION, GROUNDED IN NEUROSCIENCE AND SUBJECTIVE MEASUREMENTS.



Note: M = mental event, P = physical property concepts, m = measures of mental events, p = physical realizers of physical property concepts, ⟶ = physical causation, ⇢ = mental causation, ——— = mind-body relations, --------- = measurement relation

FIGURE 9: FEAR, ANXIETY, AND CONSCIOUSNESS



Note: See text for abbreviations

APPENDIX: GLOSSARY

Dualism (see Robinson, 2016): The theory that the mental and the physical are in some sense different kinds of thing. This is most commonly discussed in terms of the mind and the body / brain.

Eliminativism (see Ramsey, 2013): Eliminitivism is synonymous with eliminative materialism. This view holds that either some or all of the mental states posited by common-sense / folk psychology do not exist.

Emergentism (see O'Connor and Wong, 2015): If some entity can be said to "arise" out of a more fundamental entity or set of entities, yet is somehow novel or irreducible to the lower-level entity/entities, it is said to be emergent. Another way of thinking of emergence is as a process by way of which larger entities somehow arise from the interactions between smaller / simpler entities, and the larger entities exhibit properties that the smaller ones do not. As an example, consciousness is sometimes said to be an emergent property of the brain. Emergentism itself is simply the belief in emergence.

Folk psychology (see Ravenscroft, 2016): Folk psychology can be discussed in a number of different ways. In one sense, the term folk psychology refers to a set of cognitive capacities including the capacity to predict and explain behavior, among others. Folk psychology can also be used to refer to a general theory of behavior represented in the brain. In this sense then, the cognitive capacities referred to by the first meaning of folk psychology would be underpinned by the second meaning of folk psychology. Finally, folk psychology can be conceptualized as a theory of human psychology made up of the general everyday common-sense notions about the mind that laypeople generally subscribe to.

Functionalism (see Levin, 2013): Functionalism holds that what makes something a thought, pain, desire, or any other mental state is dependent only on its function. That is, the role it plays in the system it is a part of. In other words, functionalist theories claim that the identity of a mental state is given by its causal relations to (for example) behaviors, other mental states, and sensory stimulations.

Idealism (see Guyer and Horstman, 2015): Idealism can be thought of in at least two ways. Epistemological idealism is the view that human knowledge is essentially determined by human thought. Ontological idealism holds that epistemological idealism is a source of the truth because it is reality itself that is some form of thought. Within contemporary philosophical discourse however, idealism seems to be conceptualised in two different ways, a) that reality is ultimately founded on, or is in fact exclusively, something mental (e.g. the mind, spirit, reason), or b) that although one may concede to the existence of something independent of the mind, everything that we could know about mind-independent reality is so steeped in the activities of the mind, that any claim to knowledge must in some sense be self-knowledge. Idealism of form a) could be termed

metaphysical or ontological idealism, and idealism in form b) could be termed formal or epistemological idealism.

Identity thesis (see Smart, 2014): The argument that mental states and processes are identical to the states and processes of the physical brain.

Neutral Monism (see Stubenberg, 2016): Neutral monism is a variant of monism which holds that reality is neither mental nor physical. However neutral monists often also imply that reality is *both* mental and physical. Unfortunately, this tends to confuse neutral monism with other theories. Thus, more recent work (Stubenberg, 2016, and also see the present article) recommend that neutral monism should instead be taken to refer to the neither interpretation, rather than the both.

Physicalism (see Stoljar, 2016). Physicalism is closely related to a materialist substance monism view. It is the view that everything in the universe that is real is physical. To cope with contemporary counters (e.g. concerning the mind), sometimes, physicalist views instead argue that everything supervenes on the physical.

Reductionism (see van Riel and Van Gulick, 2016): To say that a given entity *a* reduces to entity *b* is ultimately to imply that *a* is nothing more than, or nothing over and above *b*. In other words, *b* is more basic than *a*, and *a* somehow depends on, or is constituted by, *b*. The ultimate expression of reductionism in a scientific context is that everything ultimately reduces to physics (e.g. social science reduces to psychology, which reduces to neuroscience, which reduces to biology, which reduces to chemistry, which reduces to physics). Obviously, such ideas are not universally popular (e.g. see Cartwright, 1999).

Substance monism (see Schaffer, 2016): The view that all things that exist are of one type. This "type" is often considered to be physical matter, but other types of substance monism would include mental, or neutral, monisms.

Supervenience (see McLaughlin and Bennett, 2014): Some property (or set of properties) supervenes on another in the case that no two things can differ on the first property/properties without also differing on the second. In other words, if *A* supervenes on *B*, there can be no difference in *A* without a difference on *B*.

The Principle of Charity (see Gutting, 2015): The Principle of Charity requires that we develop our arguments on the basis of the most favourable (i.e. most plausible and defensible) interpretation of the position of our opponents.

The Principle of Relevant Evidence (see Gutting, 2015): The Principle of Relevant evidence concerns the basis for inductive reasoning. It states that one can rely on a conclusion from an inductive argument only if one has strong reason to believe that the premises include all relevant evidence. In other words, one must be able to defend the position that there are no other facts or evidence that would be likely to alter the conclusion.