

NAMING OUR CONCERNS ABOUT NEUROSCIENCE: A REVIEW OF BENNETT AND HACKER'S
PHILOSOPHICAL FOUNDATIONS OF NEUROSCIENCE

DAVID W. SCHAAL

STANFORD UNIVERSITY

Bennett and Hacker use conceptual analysis to appraise the theoretical language of modern cognitive neuroscientists, and conclude that neuroscientific theory is largely dualistic despite the fact that neuroscientists equate mind with the operations of the brain. The central error of cognitive neuroscientists is to commit the *mereological fallacy*, the tendency to ascribe to the brain psychological concepts that only make sense when ascribed to whole animals. The authors review how the mereological fallacy is committed in theories of memory, perception, thinking, imagery, belief, consciousness, and other psychological processes studied by neuroscientists, and the consequences that fallacious reasoning have for our understanding of how the brain participates in cognition and behavior. Several behavior-analytic concepts may themselves be *nonsense* based on thorough conceptual analyses in which the criteria for *sense* and *nonsense* are found in the ways the concepts are used in ordinary language. Nevertheless, the authors' nondualistic approach and their consistent focus on behavioral criteria for the application of psychological concepts make *Philosophical Foundations of Neuroscience* an important contribution to cognitive neuroscience.

Key words: cognitive neuroscience, conceptual analysis, dualism, mereological fallacy, reductionism

For behavior analysts interested in neuroscience, the current state of neuroscientific theory is unfortunate. Despite a pleasing ring to the phrase "behavioral neuroscience," what exists instead is a thoroughly cognitive neuroscience. Researchers performing the cleverest and most inspiring neuroscience research have adopted the language of cognitivism uncritically, and often enthusiastically, so that the form of neuroscience explanations, which have the clear advantage of being supported by data collected from actual brain tissue, differs little from cognitive explanations, in which the explanatory mechanisms are inferred from observations of behavior in context. One can recognize a clear Cartesian dualism in cognitive neuroscience, even if "mind" is largely replaced by "brain," and even when the brain's activity is measured.

Bennett and Hacker recognize this fact of modern cognitive neuroscience, and criticize it thoroughly in their book, *Philosophical Foundations of Neuroscience*. They write:

For the characteristic form of explanation in contemporary cognitive neuroscience consists in ascribing psychological attributes to the brain and its parts *in order to explain* the possession of psychological attributes and the exercise (and deficiencies in the exercise) of cognitive powers by human beings. (p. 3)

That this is the form of explanation common in cognitive *psychology*, with mind (or its hypothetical components) substituted for brain, has been noted by dozens of behavioral scientists and theorists, none more consistently or effectively than Skinner. But despite generations of argumentation to the contrary, cognitive neuroscience is still largely a dualistic, reductionistic enterprise. Well-known facts about the brain, such as the homuncular organization of the sensorimotor cortex, often seem to support such a conception. The continued and often successful mission to localize cognitive function in the brain seems to answer the question "where" (as in "Where is memory stored, where is fear represented, and where is the rat's spatial map?"), and in so doing gives many neuroscientists confidence that they are on the right track.

Bennett and Hacker use conceptual analysis in the tradition of Wittgenstein to argue against these most basic assumptions of neuroscientists. Their arguments are rooted in the history of thought about the brain and mind, in extensive and scholarly reviews of the

Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Malden, MA: Blackwell Publishing.

The author thanks Drs. Tim Hackenberg and Jane Bailly for helpful discussions as this review was being written.

David Schaal is now at Accuray, Inc. Please address correspondence to David W. Schaal, Accuray, Inc., 1310 Chesapeake Terrace, Sunnyvale, California 94089 (e-mail: dschaal@accuray.com).

doi: 10.1901/jeab.2005.83-05

theoretical language of modern cognitive neuroscientists, and in careful logico-grammatical analyses of psychological concepts. Although they praise neuroscientists for their accomplishments (Bennett is a neuroscientist) and express confidence that neuroscientists will elucidate the brain activity that makes learning, thinking, remembering, imagining, perceiving, and so forth, possible, they state clearly what neuroscience *cannot* do:

What it cannot do is replace the wide range of ordinary psychological explanations of human activities in terms of reasons, intentions, purposes, goals, values, rules and conventions by neurological explanations And it cannot explain how an animal perceives or thinks by reference to the brain's, or some parts of the brain's, perceiving or thinking. For it makes no sense to ascribe such psychological attributes to anything less than the animal as a whole. It is the animal that perceives, not parts of its brain, and it is human beings who think and reason, not their brains. The brain and its activities make it possible for us—not for it—to perceive and think, to feel emotions, and to form and pursue projects. (p. 3)

This quotation expresses the theme of the book, that it is usually nonsense to ascribe to the brain psychological concepts that make sense when ascribed to whole humans (and often other animals). This explanatory tendency of neuroscientists (and cognitive psychologists) is called the *mereological fallacy*. Although occasionally it leads the authors to say things about psychology that behavior analysts would not generally agree with, the arguments against the mereological fallacy in theories of memory, perception, thinking, imagery, belief, and other psychological processes upon which the methods of neuroscience have been brought to bear will be music to the skeptical ears of most behavior analysts.

HISTORY

In the first part of the book the authors present a history of thought about the brain that begins with Aristotle (who is quoted on p. 15; “For it is surely better not to say that the soul pities, learns or thinks, but that the man does these with his soul”) and leads, not surprisingly via the mind(soul)-body dualism of Descartes, to the current form of brain-body dualism practiced by most neuroscientists.

Investigators repeatedly struggled, but persisted in incorporating the earliest neuroscientific findings (e.g., that movement can be stimulated in headless animals, and sensorimotor circuits at the level of the spinal cord are functional in decerebrate animals) into explanations that invoked the soul and animal spirits. The authors make clear the likely significance of classic neuroscientific findings, for example, the localization of language function in Broca's area and the somatotopic organization of the motor cortex, for the developing field that would be pioneered by Sherrington (1951). Despite his brilliant research and deep thought about the problem, Sherrington was unable either to explain how mind emerges from brain or to abandon mind-brain dualism.

Bennett and Hacker conclude their history with a review of the research of famed neurosurgeon Wilder Penfield. In order to remove epileptogenic tissue in the temporal cortex it was necessary for patients to be awake so that the diseased and healthy tissue could, using electrical stimulation, be delineated and healthy tissue spared. Penfield showed that patients could recall experiences in vivid detail during such stimulation, or be unable to say what they were being shown despite recognizing it, and that occasional bilateral hippocampal damage due to surgery could result in an inability to remember events that occurred to them after the surgery. Patients would express frustration at knowing what an object is but not being able to get the word for it out, or amazement that their limbs could be moved involuntarily by stimulation. After a lifetime of such work Penfield (1975) proposed a theory of mind, in which the mind reasons, decides, and understands, and directs the brain in its activity which moves the animal according to its motives. The theory differed little in form from Descartes', despite a clearer picture of the brain activities involved in behavior indicative of mind.

The authors' own view of the mind allows them to criticize dualistic theories without obvious recourse to the mereological fallacy:

The mind . . . is not a substance of any kind. Talk of the mind is merely a *façon de parler* for talk about human powers and their exercise. We say of a creature . . . that it has a mind if it has a certain range of active and passive powers of intellect and will—in particular, conceptual

powers of a language-user that make self-awareness and self-reflection possible. The idioms that involve the noun 'mind' have as their focal points thought, memory and will. And they are all readily paraphrasable into psychological expressions in which the word does not occur . . . (pp. 62–63)

Penfield's observations in the operating room provided no real support for his theory of mind-brain interaction. For example, electrical stimulation that disrupts a person's ability to name an object despite knowing what the object is does not reveal the disconnection of mind from brain, but only that stimulation disrupts some psychological functions while sparing others. Essentially the same description applies to the fascinating and important split-brain phenomenon, about which much dualistic theorizing has been generated, as the authors note frequently throughout the book.

PERCEPTION AND REPRESENTATION

The "representation" is a weed in the neuroscientific garden, and the sooner it is uprooted the better. (p. 143)

In their chapter on sensation and perception (chapter 4) Bennett and Hacker describe theories of perception that depend on the concept of representation. Modern representationalist theories hold that representations are the brain's symbolic descriptions or interpretations (as opposed to isomorphic copies) of the world, from which the brain draws inferences about what is really there. Most neuroscientists differ little from British Empiricists in their belief that we do not perceive the world but, rather, its effects on us, or the ideas it causes in us. The authors' primary argument against such theorizing is that it is nonsense: brains do not use symbols or form descriptions, human beings do, and seeing (or apprehending the world with any of the senses) is not to interpret or construct descriptions of anything. Furthermore,

To say that the mind has 'access' to the 'internal representation' produced by the brain is no less mysterious than the Cartesian claim that the mind has access to an image on the pineal gland. Moreover, it is altogether obscure how the mind's having access to putative neural *descriptions* will enable the person *to see*. And if [David] Marr were to

insist (rightly) that it is the person, not the mind, that sees, how is the transition from the presence of an encoded 3-D model description in the brain to the experience of seeing what is before one's eyes to be explained? To be sure, *that* is not an empirical problem . . . It is the product of a conceptual confusion, and what it needs is disentangling. (p. 147)

This argument is reminiscent of Skinner's (1969) concerns about the brain's role in seeing:

Suppose someone were to coat the occipital lobes of the brain with a special photographic emulsion which, when developed, yielded a reasonable copy of a current visual stimulus. In many quarters this would be regarded as a triumph in the physiology of vision. Yet nothing could be more disastrous, for we should have to start all over again and ask how the organism sees a picture in its occipital cortex, and we should now have much less of the brain available in which to seek an answer. (p. 232)

Most behaviorists will appreciate the authors' critique of "representation" in neuroscience, having made many of the same points about cognitive psychology over the years, but it remains unclear what to make of the neural phenomena that neuroscientists refer to by the term. As neuroscience methods improve (in particular, *in vivo* real-time imaging and electrophysiology) we appear closer and closer to having the neural referent of the representation. In other words, the representation seems to be less an inference from behavioral data and more an observed neural fact. How do we deal with this? What is "representation" in this context?

Bennett and Hacker do not deny a sense in which there are representations in the brain. Activity in parts of the brain may serve as correlates of features of the object perceived, may *represent* features of the object in the simple sense that changes in their activity are caused by those features. In this limited sense, however, the concept has none of the explanatory value that cognitive neuroscientists require of it, and as such the concept can be dispensed of without loss. If the brain does not use sense data to construct internal representations, if those representations do not function as symbolic descriptions of the perceived world, then a preserved, limited concept of representation is of no value and the conception of perception itself needs revision.

The authors' conception of perception is consistent with a behavioral one, indicated by claims such as, "Perceiving is an epistemic relation between a perceiver and an object perceived" (p. 128) and "Possession of a sense-faculty is manifest in behaviour" (p. 127). Behaviors that are appropriate to things seen (finding one's keys or calling a red ball "red") are the logical criteria for saying that one sees. Brain events may be correlated with perceived objects and participate causally in the person's seeing, but the seeing is always a function of persons, not their brains. Given this, the authors characterize the task of a neuroscience of seeing:

Which neuronal groups must simultaneously be active in order to achieve optimal vision, what form that activity may take, and how it is connected with other parts of the brain that are causally implicated in cognition, recognition and action, as well as in co-ordination of sight and movement, are what needs to be investigated by neuroscientists. (p. 142)

In other words, discovering how the brain participates in the relation between the animal and its environment called perceiving, how it makes perceiving possible, is what a neuroscience is for. Much the same could be said of the other psychological functions; the brain is a critical participant in orderly relations between organisms and their environments. The brain's role is not to execute psychological functions (the brain does not make decisions), or contain them (the brain does not have images), or acquire them (the brain does not learn); people do these things, and in so doing depend on their brains.

COGNITION AND COGITATION

In chapters 5 and 6 the authors examine the theories and approaches of modern neuroscience to psychological powers such as knowing, remembering, thinking, and imagining, and in so doing, I suspect, name most of the concerns of behavior analysts about neuroscience. *Knowing* is conceived by the authors, not as a state of the brain, but as "ability-like"; "For language-using creatures such as ourselves, to know where, when, who, what, whether, and how . . . is, among other things, to be able to answer these questions" (p. 149). This conception follows from an analysis of the ordinary

contexts in which the words *know*, *knowing*, and *knowledge* occur. Criteria for saying that someone knows something do not include references to the state of their brain, but rather to the behavior that is indicative of their knowing. Knowing is not an activity of the brain but of human beings, and knowledge is not contained in the brain but in books and computers, and is possessed by human beings, but not by their brains. It makes no sense and explains nothing to divide the brain up into bits that contain different kinds of knowledge and know different sorts of things, because the brain does not contain knowledge or know anything. A split-brain patient differs from someone with an intact corpus callosum not in the sense that knowledge one side of the brain possesses is not shared with the other side (so the right brain knows, for example, which object should be selected with the left hand but the left brain does not know the name for the object). In chapter 14 the authors describe the split-brain phenomenon and offer an explanation that is faithful to the findings without attributing acts of seeing, knowing, or interpreting to the brain:

. . . the general form of the explanation is that severing the corpus callosum deprives human beings of the capacity to exercise normally coordinated functions. And that in turn is to be explained in terms of the disconnection of neural groups that are causally implicated in the exercise of the relevant capacities. (p. 393)

A similar "form of explanation" applies to the phenomenon of blind-sight, a condition caused by damage to the right occipital lobe (see also chapter 14). Patients are unable to see in the sense of naming objects before them or even saying that there are objects before them, yet they are able to identify them in forced-choice tasks or to avoid objects while moving through rooms. Although neuroscientists have explained blind-sight by saying that the brain's ability to sense visual stimuli is separate from its ability to monitor its own sensations, blind-sight shows instead that some kinds of behavior indicative of seeing can be dissociated from other kinds of behavior indicative of seeing. Donahoe and Palmer (1994, pp. 23–24) reach a similar conclusion; worded more generally, environment-behavior relations that usually "go together" may be separated in neuroscience experiments in

a way that allows inferences to be drawn about the brain areas that participate in them.

At this point in the review it should not be surprising to learn that that Bennett and Hacker take a dim view of most modern memory theory. Conceptions of memory as “stored representations of antecedent experience,” as “neural traces” and “encoded information,” are deficient. The concept of “storage” is so far removed from its counterpart in ordinary language that it does more harm than good. And to conceive of memory as the storage of previous experiences leaves out much of what counts as memory. For example, I remember that the square root of 81 is 9, but I have no recollection of the occasion upon which I learned this, and even if I did, my answer “9” to the question “What is the square root of 81?” does not require that the learning occasion be brought to mind. The authors describe the phenomena that fall under the heading of “memory” in this way:

Memory is the retention of knowledge previously acquired. It is an ability that may be exercised in indefinitely many forms: for example, in *saying* what one remembers, affirming *that* one remembers it when asked, not *saying* anything but *thinking* about what is remembered, neither saying nor thinking anything but acting on what one remembers in any of indefinitely many ways, recognizing something or someone, and so forth. It is very tempting to think that the diverse forms in which remembering something may be manifest are all due to the fact that what is remembered is *recorded* and *stored* in the brain. But that is a nonsense. (p. 170)

This does not mean that a neuroscience of memory is not necessary, but it does recast the problem fundamentally into a search for the neural preconditions and concomitants of remembering. An outcome of the authors’ reasoning is a reconception of modern “memory systems” approaches, according to which different types of memory are “stored” in different locations in the brain. For example, episodic memory (which concerns remembering of the context and sequence of events one has experienced) has repeatedly been shown to depend on an intact hippocampus, and procedures that result in episodic memory alter the firing pattern of populations of hippocampal neurons (see Eichenbaum & Fortin, 2005). These facts lead some to

conclude that the hippocampus is critically involved in the storage of episodic memories. About this conclusion the authors remark: “It may be that the retention of certain synaptic connections and the creation of certain recurrent firing patterns are a necessary condition for one to be able to recall something—but that is all” (p. 170). The referent of the term, “episodic memory,” is the behavior-in-context that indicates that an animal has retained what it has learned, not the neural events that make the retention possible (neural events which it is the task of neuroscience to discover).

Despite their critique of memory, and conception of it that clearly is more congenial to a behavioral approach, the authors propose restrictions on the concept that seem unnecessary. Neuroscientists have uncovered, to a great extent, the neural mechanisms underlying simple forms of learning in rodents and invertebrates, and although many of us (e.g., Villarreal & Steinmetz, 2005) would allow such phenomena to be classed as instances of memory (in the sense that lasting changes in behavior are wrought by experience), the authors say that “most of it is not research on memory in any sense of the word . . .” (p. 156). I see their point, but it seems likely that the tissue changes observed in such simple models reflect, at least in part, the changes that underlie the phenomena the authors classify as true memory (the authors admit that this is a possibility). Animal behavior researchers may be more inclined to see similarities in diverse sorts of behavior from different species—similarities that are the basis for beliefs that animal behavior models aspects of human behavior. I may be among the neuroscientists that the authors criticize in the sense that I see no harm in allowing the concept of memory to be applied in these simple cases (although, as a behaviorist, I see no point in doing so).

In chapter 6 the authors consider “cognition,” the human powers of belief, thought, and imagination. As is true of all the psychological powers confronted in the book, these are conceived as abilities and capacities of humans, not of their brains or parts of their brains. A consideration of the varieties of powers, such as thinking, reveals both that neuroscientists tend to study only a small portion of the phenomena that count as

thinking, and that it is nonsensical to apply the concept to the brain. Neither does it make sense to think of the brain as the locus of thought. Instead, my thoughts occur in my office, or my car, or wherever I am when I am having a thought. "For a thought is just what is expressible by an utterance or other symbolic representation," the authors state (p. 180). One may debate with them whether the brain should be conceived as the organ of thought (they say no), but this does not lessen the common-sense thrust of their treatment of these concepts.

CONSCIOUSNESS

A large portion of the book is devoted to the concept of consciousness. The authors present several quotations from famous neuroscientists and philosophers that indicate clearly that consciousness is conceived by many as the great mystery left to be solved, one which may never be solved. Its mystery inheres in its supposed privacy, a characteristic the authors disdain:

On this widely shared conception, our alleged ignorance is explained by reference to the thought that each person has privileged access to his own consciousness, but not to the consciousness of others. So consciousness is not a publicly observable, but a privately observable, phenomenon and, in this respect, unlike the phenomena typically studied by the sciences . . . this conception of privacy is confused. (p. 241)

Because the criteria for determining whether or not a person is conscious in a given instance are behavioral, consciousness is not private in the sense of being inaccessible (for example, in the simplest case, it is usually clear when one becomes conscious in the sense of waking up, even if occasionally people pretend to be asleep).

The mysteriousness of consciousness also stems from its association with experience, as in *conscious experience*. Although it is possible, according to the authors, to be conscious of an experience one is having or to have an experience while one is conscious, conceptual danger arises when *conscious* is thought to be a property of experience (i.e., when it is the experience that is conscious and not the person). This conception goes beyond consciousness in more simple senses (conscious as

opposed to asleep, or *conscious of* in the sense of having one's attention caught and held by something), into a mysterious realm that reveals the dualism (what the authors call "crypto-Cartesianism") that still grips cognitive neuroscientists. This brings us to the "qualitative character of experience," a view of consciousness originated by the philosopher Thomas Nagel (1974) and adopted by many neuroscientists. The authors quote Nagel:

. . . the fact that an organism has conscious experience *at all* means, basically, that there is something it is like to *be* that organism . . . fundamentally an organism has conscious mental states if and only if there is something that it is like to *be* that organism—something it is like *for* the organism. (p. 272)

The authors write, "For every 'conscious experience' or 'conscious mental state' there is something which it is like for the subject to have it or to be in it" (p. 273). It is not simply being conscious of a thing or state of affairs, but to be conscious of the quality of the thing or state of affairs (the "qualitative feel"), that is the sort of consciousness that many neuroscientists equate with a great mystery, an uncharted frontier. This concept is central to the active discussion of the special characteristics of "zombies," which are creatures exactly like humans in every respect except for not having *qualia*, or the qualities of conscious experience (for an introduction to zombies see Kirk, 2003).

How the authors treat the concept of *qualia* typifies the philosophical rigor with which they approach all the psychological concepts in the book. Despite the fact that they declare it incoherent, they take it seriously in order to determine what neuroscientists mean by it. Much of their discussion concerns the strange phrase "something which it is like." One may ask someone "what it is like" to have an experience, but here what we seem to be requesting is an appraisal of the experience, one's attitude about it, as in "I found the experience unpleasant." But having an attitude about something applies only to a limited range of that which we are conscious of, and besides, this is not what neuroscientists intend when they discuss conscious experience or *qualia* in this manner. Applied to individuals, we may legitimately ask, for example, "What is

it like to be a doctor?” and expect to get an appraisal of a doctor’s activities and lifestyle and attitudes, but one interested in qualia would ask, “What is it like for a doctor to be a doctor?” and this is a more difficult question to grasp, one which the authors eventually declare to be illicit. Ultimately, the notion of qualia reveals that, for many neuroscientists, there exists a fundamentally subjective, private, difficult or impossible-to-describe realm of personal experience that is consciousness. It is this conceptual muddle that naturally leads neuroscientists to view consciousness as a deep mystery.

A careful consideration of the logical and grammatical characteristics of consciousness suggests that it is not a single thing, but instead is indicated by many human and animal behaviors and abilities:

We attribute consciousness to a creature on the grounds of its behaviour in the circumstances of its life, not on the grounds of its possessing private qualia or movies-in-the-brain The behaviour that warrants the attribution of one form of consciousness differs from that which warrants attribution of another there is no sharp divide in nature between creatures to which it makes sense and creatures to which it makes no sense to ascribe consciousness or experience in one or other of their many forms. Rather . . . as their nervous systems, perceptual organs and brains become more evolved, more and more forms of apprehension of, and response and reaction to, their environment become possible there is a gradual evolution of more and more complex forms of sensitivity to the environment and more and more complex forms of response. (pp. 303–304)

The authors’ conception removes some of the mystery from the notion of consciousness, and suggests that aspects of consciousness can be profitably studied in animals other than humans. According to their view, a neuroscience of consciousness would not necessarily be a unified field (consciousness is not a unified phenomenon), but would have to deal separately with consciousness in all its forms.

Finally, the authors conceive consciousness as something we become, but not something we do or can be trained to become. “One can be good at learning, discovering, detecting or finding out certain things, but one cannot be good at becoming conscious of things”

(p. 256), and “becoming conscious, becoming aware, etc., are not things we do, let alone actions we perform” (p. 257). How this can be true and there still be a category of consciousness the authors call *dispositional consciousness* is not clear. Dispositional consciousness is a general tendency to be conscious of certain things—money-conscious, for example. Such a generalized tendency is indicated by various sorts of behavior—money-conscious people are likely to save their money, spend it carefully, talk about it and think about it more than others, and so forth. Such a tendency almost certainly is learned, and therefore one can be “better” or “worse” at it depending on one’s experience, if “better” and “worse” refer to a greater or lesser probability of behaving in ways consistent with the disposition. So the authors’ assertion that consciousness is not something we can become “good at” may be argued with, both in its dispositional sense and in its *occurrent transitive* sense (a current consciousness of some thing or state of affairs). I may not become conscious of the subtle French horn part in a piece of music until after I have read about the composer’s penchant for using the French horn in subtle ways—has my learning not enhanced my ability to be conscious of the French horn in the composer’s music? More broadly, is there no sense in which the common Californian pastime of “expanding” or “developing” consciousness is true?

NOT BEHAVIORISM

Although I suspect their arguments will be greatly appreciated by readers of this journal, the authors are not behaviorists. (This may be strategically beneficial; in a postcognitive-revolution world it may be best for nondualistic, nonreductionistic approaches to neuroscience not to be also identifiable as behavioristic, lest readers dismiss them after considering the source.) The authors perform conceptual analyses *à la* Wittgenstein, in which the meanings of concepts as they evolved naturally provide the basis for judgments of “sense” or “nonsense” when the concepts are used in neuroscience. Thus, for example, the notion of memory storage is nonsense because (among other reasons) “. . . even if there were such a ‘record,’ it would not be *available* to a person in the sense in which his diary or

photograph album is available to him” (p. 164).

Behavior analysts will usually find the results of the authors’ analyses congenial because of their consistent application of behavioral criteria for the use of psychological concepts. But the application of criteria handed down from everyday language to determine whether a concept makes sense in a scientific context renders as nonsense several conceptions of radical behaviorists. For example, although their ideas about the notion of privacy are compelling and largely consistent with a behavioral approach (e.g., Skinner, 1945), the authors have no place for a special notion of privacy. For them “private” is what happens when a person is alone, or is a feeling or image a person has that the person does not report (see chapter 3 for a description of Wittgenstein’s arguments on the topic). As another example, some behaviorists conceive of “imagination” as “seeing in the absence of the thing seen” (Skinner, 1974, p. 91). Skinner wrote: “when a person sees a person or place in his imagination, he may simply be doing what he does in the presence of the person or place.” But for the authors, “*Seeing* is not something done.”

There are senses in which we would agree with Bennett and Hacker on these issues, and even on their thoroughly maintained distinction between mental and behavioral events, largely because they are thorough in their assertion that the criteria by which we identify mental events should be behavioral ones. In their excellent chapter on reductionism, however (chapter 13; readers should consider reading this chapter first), the authors declare that there are no psychological laws:

Not only are there no bridge principles allowing any form of ontological reduction of psychological attributes to neural configurations, but it is far from evident that there is anything that can be dignified by the name of *psychological laws* of human action (p. 362)

We can “explain” a person’s behavior by citing his reasons for behaving in such-and-such a way in such-and-such a context, but that does not constitute consistency with psychological laws, of which there are none. This argument made no sense to me, and seemed, relative to the goals of the book, to be beside the point. For the authors it seems that the fact

that humans behave differently in similar contexts negates the possibility of psychological laws, despite the fact that it is the mission of many of us to determine what historical factors led to the different behavior. We may agree with them that a neuroreductionistic explanation of a bit of human behavior may shed little light on why the behavior occurred, without also asserting that the behavior is not indicative of the operation of psychological laws.

CONCLUSIONS

Orderly relations between neural activity and behavior seem to provide support for cognitive explanations of the behavior. The form and content of neuroscientists’ cognitive explanations are guided (at least in part) by neuroscientists’ contact with the observed behavior-in-context, in which brain activity participates (more or less critically). Neuroscientists looking for brain activity correlated with the behavior-in-context will find it, and that will support their cognitive explanations and their overall dualistic conception. This book suggests alternatives to those explanations. The adoption of these alternatives could result in greater focus on the environment-behavior relations that are at the heart of behavior analysis (and psychology, or so we would have it), and perhaps draw behaviorists into the field in greater number, resulting ultimately in a more truly behavioral neuroscience. As one who believes that neuroscience will be a critical part of the evolution of basic and applied behavior analysis, then, I highly recommend this book. The contribution of behavior analysis to neuroscience may depend on a successful defeat of the mereological fallacy in cognitive neuroscience.

Although something like Bennett and Hacker’s view of the language of neuroscience may be necessary for a conceptual rapprochement between behavioral events and neural events, it is entirely possible that the real conceptual puzzles are of a different sort. Far more difficult to achieve, I believe, will be an understanding of the fundamental *nestedness* of the brain, the rest of the body, and the person in the world, each entity executing processes that overlap and turn back on themselves and each other in time and space. The firing of a neuron in the lateral intraparietal area may be critical to the execution of

a choice response that is reflective of recent relative reinforcement rates (see, e.g., Corrado, Sugrue, Seung, & Newsome, 2005; Lau & Glimcher, 2005), but the individual neuron's firing only has meaning when it is part of an integrated neuronal circuit (in this case, part of the oculomotor circuit), the activity of which only has meaning relative to the current environmental-behavioral context (the events arranged in their concurrent schedule procedures), which itself only has meaning relative to previously experienced environmental-behavioral contexts (the extensive training the animals received). I suspect a sufficient understanding of how the brain participates in behavior will depend on an ability to refer simultaneously to events at multiple levels of integration and at multiple time frames, including—most importantly from the perspective of behavior analysts—the animal's history. Neural causation will not be able to replace mnemonic causation. As described by Bertrand Russell in his 1921 lectures published as *The Analysis of Mind*, *mnemonic causation* requires that an explanation of a behavior in a current setting include references to

. . . past occurrences in the history of the organism as part of the causes of the present response. I do not mean merely—what would always be the case—that past occurrences are part of a CHAIN of causes leading to the present event. I mean that, in attempting to state the PROXIMATE cause of the present event, some past event or events must be included, unless we take refuge in hypothetical modifications of brain structure. For example: you smell peat-smoke, and you recall some occasion when you smelt it before. The cause of your recollection, so far as hitherto observable phenomena are concerned, consists both of the peat smoke (present stimulus) and of the former occasion (past experience). (p. 57)

Russell held that mnemonic causation was necessary “unless we take refuge in hypothetical modifications of brain structure,” and although it may be thought that mnemonic causation could be dispensed with when the modifications of brain structure are no longer hypothetical, such is not the case. Alterations in brain structure and function may allow the past to govern an animal's current behavior (i.e., may mediate between prior experience and current behavior), but such alterations are themselves mnemonic phenomena; their mean-

ing can be fully understood only in light of the animal's history. The apparent unwillingness of neuroscientists to allow references to the animal's past to play key roles in their explanations of behavior may be their more important conceptual difficulty.

Finally, metaphors and analogies are always hung on the edges of scientific understanding, and thorough adoption of the practices suggested in this book will not change that. I am led by my colleagues in stroke research to refer to intra- and extracellular events that lead to cell death or survival following cerebral ischemia as comprising a stream or pathway, in which the events lie upstream or down the path from each other. For example, ischemia leads to oxidative stress in mitochondria that, downstream, causes cytochrome C to be released into the cytoplasm that, downstream, activates caspases that result in DNA damage and cell death by apoptosis. We use words like *stream* and *pathway* in this context to capture in a simple way their sequential nature while remaining somewhat noncommittal about their casual relations, but we know that they are deficient, that they misrepresent the complexity of the situation. They are certainly an improvement over referring to these events as links in a causal chain, however, because they capture a little better the fact that the events are arrayed probabilistically. But the banks of the metaphorical stream and the fact that it flows in one direction are constraints of the model that certainly will be violated by the phenomenon itself. Maybe *current*, as in an *ocean current*, is more apt in this context because the boundaries are much wider and the flow less insistent. On the other hand, maybe the simpler *stream* is better. The point is that it may be the ability of metaphors and analogies to help researchers accomplish their theoretical goals, and not how well they stand up to connective analysis relative to their conventional counterparts, that is the better basis for approving or disapproving of them.

REFERENCES

- Corrado, G. S., Sugrue, L. P., Seung, H. S., & Newsome, W. T. (2005). Linear-nonlinear-Poisson models of primate choice dynamics. *Journal of the Experimental Analysis of Behavior*, *84*, 581–617.
- Donahoe, J. W., & Palmer, D. C. (1994). *Learning and complex behavior*. Needham Heights, MA: Allyn & Bacon.

- Eichenbaum, H., & Fortin, N. J. (2005). Bridging the gap between brain and behavior: Cognitive and neural mechanisms of episodic memory. *Journal of the Experimental Analysis of Behavior, 84*, 619–629.
- Kirk, R. (2003). Zombies. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Available from <http://plato.stanford.edu/archives/fall2003/entries/zombies/>
- Lau, B., & Glimcher, P. W. (2005). Dynamic response-by-response models of matching behavior in rhesus monkeys. *Journal of the Experimental Analysis of Behavior, 84*, 555–579.
- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review, 83*, 435–450.
- Penfield, W. (1975). *The mystery of the mind: A critical study of consciousness and the human brain*. Princeton, NJ: Princeton University Press.
- Russell, B. (1921). *The analysis of mind*. London: George Allen & Unwin.
- Sherrington, C. S. (1951). *Man on his nature*. (2nd ed.). Cambridge, England: Cambridge University Press.
- Skinner, B. F. (1945). The operational analysis of psychological terms. *Psychological Review, 52*, 270–277.
- Skinner, B. F. (1969). *Contingencies of reinforcement: A theoretical analysis*. New York: Appleton-Century-Crofts.
- Skinner, B. F. (1974). *About behaviorism*. New York: Alfred A. Knopf.
- Villarreal, R. P., & Steinmetz, J. E. (2005). Neuroscience and learning: Lessons from studying the involvement of a region of cerebellar cortex in eyeblink classical conditioning. *Journal of the Experimental Analysis of Behavior, 84*, 631–652.