

PROBING THE LIMITS OF MIND AND BRAIN

Dave Suarez
Undergraduate Student
Simon Fraser University
E-mail: dsuarez@sfu.ca

*The Brain is wider than the Sky
For put them side by side
The one the other will contain
With ease and You beside
- Emily Dickinson (1896)*

1. INTRODUCTION

This essay has two objectives: first, to show the relevance of cognitive neuroscience to philosophy, and second, to propose an area of research that can be addressed using philosophical and neuroscientific methods.

Regarding the first objective, Stone & Davies (1993) have previously argued that the results of cognitive neuropsychology have implications for theory in the philosophy of mind. Their appeal to relevance is limited to the functional organization of mind, as revealed by cognitive neuropsychological data. As Stone & Davies see it, “cognitive neuropsychology reckons the *neurophysiological* details of patients to be more or less irrelevant” (emphasis mine). While I believe that this may be true of investigations in cognitive *neuropsychology*, I believe that cognitive *neuroscience*, which takes such neurophysiological details into account, can also be shown to be relevant to philosophical discussion.

With respect to the second objective, I hope to show that aspects of metaphysics can inform, and be informed, by data from cognitive neuroscience. P. F. Strawson (1959) describes the field of *descriptive metaphysics* as “content to describe the actual structure of our thought about the world.” It is becoming increasingly apparent through work in the fields of cognitive neuropsychology and cognitive neuroscience that thought is constrained in some way by brain function. I do not wish to argue here that mental states are identical to, or even reducible to, brain states, but it does seem to be an inescapable fact that there is some kind of a correlation between the two. That a correlation exists, will suffice for the moment as an assumption from

which to work. It may prove to be an incorrect assumption, but demonstrating the independence of thought from brain function remains an open challenge.

In a later section, I will show how Strawson's account of persons (1959) parallels empirical findings in cognitive neuroscience regarding 'theory of mind'. I suggest that this kind of parallelism between philosophy and neuroscience reveals a pattern of convergent results regarding the limitations of mind. Philosophy can provide a map of "the bounds of reason" in terms of conceptual structure, and cognitive neuroscience can show us how these boundaries come to be as they are.

2. THE RELEVANCE OF COGNITIVE NEUROSCIENCE TO PHILOSOPHY

In giving an argument for the relevance of cognitive *neuropsychology* to philosophers, Stone & Davies (1993) stop short of admitting the relevance of "neurophysiological details". This section outlines their argument, and extends it to include such details, bringing cognitive neuroscience into the picture

2.1. FIRST, COGNITIVE NEUROPSYCHOLOGY AND THE PHILOSOPHY OF MIND

Stone & Davies (1993) believe that cognitive neuropsychology can be of interest to philosophers in two ways. First, "philosophers can take a metatheoretical interest in the assumptions and arguments of cognitive neuropsychology." In taking up this role, philosophers give clarity and direction to cognitive neuropsychology by building theoretical foundations for methods, and helping to shape cognitive neuropsychological concepts through criticism. Metatheoretical philosophers are outside observers, commenting on the methods and concepts of an empirical field.

Second, cognitive neuropsychology can play an auxiliary role to philosophy. The investigation of certain philosophical problems (particularly in philosophy of mind) can be guided by neuropsychological data. Philosophers can, in some cases, take empirical data and use it to shed light on their own problem of interest. This second intersection of the fields of philosophy and neuropsychology will be one of the concerns of this essay.

But how can an empirical discipline like cognitive neuropsychology inform a discipline that often deals in *a priori* truth, as philosophy does? It could be argued that philosophy deals in *Canadian Undergraduate Journal of Cognitive Science, Spring 2002 Issue*
<http://www.sfu.ca/cognitive-science/journal/>

a different kind of explanation than cognitive neuropsychology. On this understanding of philosophy, empirical data would be, for the most part, irrelevant. However, Stone & Davies (1993) show that a cognitive neuropsychological explanation of prosopagnosia (a specific disability of face recognition) could provide an *alternative* to certain philosophical explanations. If their comparison of the explanations is legitimate -- if, in their fully extended interpretation, the explanations do conflict -- then we know that at least in some cases, that there is competition between neuropsychological and philosophical theory. Competition would imply that both accounts do, in fact, operate on the same level, and address the same (or at least similar) questions. For Stone & Davies' argument, the truth or falsity of either explanation is irrelevant: "[t]he possibility that the explanations are in competition is all we need to make our point."

This kind of competition can be put to good use in the realm of philosophy, particularly when cognitive neuropsychology, to use Stone & Davies' phrase, "[makes] good on failures of imagination." Empirical data can present counterexamples to what are taken to be *a priori* facts, and may reveal philosophical errors. This is not to say that all disagreements entail changes in philosophy; instead such disagreements can lead to renewed discussion.

Stone & Davies sum up the interaction of philosophy and cognitive neuropsychology beautifully when they say:

Often philosophical theory uncovers necessary conditions for the application of personal level folk psychological properties. Subpersonal level cognitive psychology tells whether, and if so how, those necessary conditions are met.

2.2. ENTER COGNITIVE NEUROSCIENCE

The interactive relationship of cognitive neuropsychology with philosophy is what opens the door for cognitive neuroscience to enter into philosophical debate. Stone & Davies (1993) admit that "the claim that neurophysiological details are not especially relevant is usually pragmatically grounded." Stone & Davies quote cognitive scientist Tim Shallice, who wrote in 1988, "[t]o hope for an advance in theories of the functional organisation of cognition by paying special attention to issues of localisation is not, *at present*, a promising strategy" (emphasis mine). While Shallice's statement may have been true when it was written, the 'unpromising strategy' of localisation has since been pursued, and it now one of the core methods of cognitive neuroscience. I would argue that relating neurophysiology to the functional structure of cognition has become somewhat easier since the writing of Stone & Davies' paper. Stone & Davies write

that, “virtually all cognitive neuropsychologists agree that . . . psychological theories are constrained from below by the facts of neurophysiology.” It is the details of these constraints that cognitive neuroscience is able to contribute.

Cognitive neuroscience augments the single case study methodology characteristic of cognitive neuropsychology with neuroimaging studies and computational modelling. With these additions, cognitive neuroscience is able to give a fuller picture of the neurological structure that underlies the functional architecture of cognition. It is an assumption of cognitive neuroscience that the functional components of cognition are realized by neural mechanisms. This can be (but is not always) differentiated from the related assumption that mind is reducible to brain function. It is not essential to cognitive neuroscience that the entirety of mental life be reducible to brain function; it is only essential that some systematic relationship exist between the mental world and the neurophysiological world for some set of mental processes. If mental processes turned out to be something entirely different from brain function, but were still inextricably bound up with brain activity, cognitive neuroscience could still proceed. At present, most of the research in cognitive neuroscience is focussed on unconscious (subpersonal) processes, which distances the field from some of the problems that plague reductionists, particularly those concerning subjective experience (e.g., qualia). The lingering issue of whether unconscious processes are correctly called ‘mental’ processes is discussed in Stone & Davies’ paper. *Contra Wittgenstein*, their conclusion is that so long as what is explained by unconscious processes comes into contact with the accounts given by philosophy, what is ‘mental,’ is merely a terminological question, and not a methodological one.

Cognitive neuroscience assumes that the operational characteristics of functional modules posited by cognitive neuropsychologists can be determined, in part, by their physical realization. In other words, how a particular mental process works can be revealed by the way that the mental process is instantiated in the brain. This is not to say that a mental process could not be instantiated differently; only that its *actual* instantiation yields clues about its function. Functional modules can be viewed in terms of the computational paradigm: functions are programs that use neural mechanisms as hardware. The kinds of functional module that one could expect to find are limited by the available neural ‘hardware’. Furthermore, certain identified functional modules are localisable to particular brain areas. These areas, in turn, appear to have an organization that is tailored specifically to the operation of that particular function. By observing neurophysiological structure and activity in restricted areas of the brain, it may be

possible to (a) discover new functional modules, and (b) refine our understanding of the internal workings of the functional modules already discovered.

What do these kinds of discoveries mean to philosophy? Stone & Davies (1993) say that,

Cognitive neuropsychology provides us with a much more fine-grained account of the mental than that given us by those pre-theoretical intuitions upon which many of our philosophical claims are (inevitably) based.

Cognitive neuroscience extends the level of specificity further, allowing functional modules to be described in terms of neurological, and perhaps even computational, form. A taxonomy of functional modules would enable more precise theories of mental process to be compared to philosophical theories, hopefully enhancing the precision of the philosophy. As Stone & Davies say, “[t]here is a tendency in philosophy to think in terms of large, poorly differentiated categories.” This tendency might be curtailed by input from the cognitive brain sciences.

3. HOW BRAINS BUILD BORDERS

3.1. STRAWSON ON PERSONS

Strawson (1959) poses two questions concerning personhood. The first question is: “why are one’s states of consciousness ascribed to anything at all?” The second question is: “why are [states of consciousness] ascribed to the very same thing as certain corporeal characteristics, a certain physical situation, &c.?” Strawson believes that the answer to both questions lies in the logical primitiveness of our concept of persons.

Strawson believes that the answer to the second question is simpler, so he gives it first. States of consciousness are only ascribed to persons, and it turns out that persons are also the kind of thing to which it is appropriate to ascribe physical predicates. Strawson believes that attempts at further decomposition are inevitably fruitless, because persons are the most basic concept to which psychological predicates are applied. The contents of our ontology do not include an entity that can be the subject of psychological predicates but not physical predicates. Such an entity, a pure ego, could not be a concept that we possess.

The answer to first question is a little more tricky. Strawson appeals to the logical form of predicate application, which requires that a predicate be contingently applicable to more than one individual. A predicate which is only applicable to one target, may as well apply to none at all: if it is true that the target possesses said predicate, it is a tautology to state that the target possesses it. In the case that the target does not possess the predicate, the predicate could have no place in any discussion, as it is never possessed by anything. It is thus a necessary condition of applying psychological predicates to ourselves that we are first able to apply psychological predicates to others. To apply psychological predicates to others, we must have some way to individuate the targets of psychological predicates. The only targets for which individuation is possible are persons, who have physical, as well as mental, traits. So the answer to the first question is that mental states are ascribed because our primitive concept of persons allows them to be applied to others, as well as to ourselves.

If it holds water, Strawson's argument has implications for psychology. An inability to ascribe psychological predicates to others, would engender an inability to ascribe personhood to oneself and to others. The concept of a person (in the Strawsonian sense) would effectively disappear from the ontology of someone with this disability. This has been borne out by psychological research on 'theory of mind'.

3.2. PERSONHOOD AND 'THEORY OF MIND'

'Theory of mind' is the psychological term for our ability to ascribe others with beliefs and desires, i.e., the ability to ascribe others with psychological predicates. People with autism display behaviour that is explainable in terms of a non-functioning theory of mind. In experimental studies, autistics have been shown to perform poorly on tasks that involve reasoning based on the attribution of mental states, while performing at a normal level on tasks that involve other kinds of inference. Autistics do significantly worse than healthy controls on what is called a 'false belief' task. A typical false belief task pairs a story with a question about what one of the characters in the story believes to be the case.

An example story (from Sabbagh & Taylor, 2000):

Ben put a folder and a clipboard on his desk.
 His friend Maggie noticed that he had lots of work to do.
 Then, Maggie went out for a coffee.
 While Maggie was gone, Ben moved the clipboard.

Ben put the clipboard on the bookshelf.
He left the folder on his desk.

This is paired with the question:

According to Maggie, where is the clipboard?

The correct answer is, of course, that Maggie (mistakenly) believes that the clipboard is on the desk. We know this because she was not in the room when Ben moved the clipboard, and could not have knowledge of the change. If autistics are not able to attribute others with any beliefs, mistaken or not, it would explain their difficulty with this kind of reasoning.

This can be contrasted with ‘false photograph’ tasks. Autistics have no trouble understanding that states of the world can be represented by photographs. Answering questions about stories in which photos represent counterfactual states of the world, is not a problem for autistics. They are able to perform at the same level on these tasks as everyone else.

To say that autistics live in a solipsistic world understates the case: not only are autistics cut off from understanding the mental lives of others, they are unable to recognize their own.

3.3. DRAWING THE BOUNDS OF REASON

Deficits in ‘theory of mind’ reasoning co-occur with pathological conditions other than autism. Damage to particular brain areas caused by stroke, trauma, or congenital defect has been observed to induce deficits similar to those found in autistics. Evidence from these lesion patients, and from functional neuroimaging studies on healthy subjects performing theory of mind tasks, indicates that theory of mind utilizes a dedicated neural substrate, and is a localisable neurophysiological module. The invariance of the development of theory of mind in children points towards a genetically pre-determined component to this module.

If theory of mind turns out to be innate, it could be seen as empirical evidence for the primitiveness of the concept of person. The functional specifications of what such a module must be like are laid out in Strawson’s explication of persons: such a module must allow us to attribute states of consciousness to a single entity to which we also attribute physical characteristics, and states of consciousness must also be attributable to others on the basis of behavioural criteria. The theory of mind module will be uncovered as fulfilling these specifications or not. If it does

not, then the issue of how we conceptualize personhood can be reopened and re-examined in light of the actual operational characteristics of the theory of mind module.

I believe, however, that the theory of mind module will fulfill Strawson's specifications for its operation. This would allow for a different kind of interaction between philosophy and cognitive neuroscience; one in which *a priori* considerations of ontological structure are found to be confirmed or disconfirmed by the cognitive brain sciences.

The bounds of human experience are amenable to investigation by both the cognitive brain sciences, and by philosophy. Philosophy is able to define the *a priori* limitations of mind, leaving the cognitive brain sciences to explain how the limitations could be contingently realized in the brain. While this kind of research is based squarely in the materialist camp, it is important to remember that the method described points only to the *possibility* of such an identity; i.e., *if* it were the case that mind were the same as brain, then the limitations of mind could be explained by some neurological, or computational, constraint on brain function.

ACKNOWLEDGEMENTS

I wish to thank Kirstie Laird for her support throughout the various stages in the development of this paper. I would also like to thank the referees for their helpful comments and criticism.

REFERENCES

- Stone, T., Davies, M. (1993). Cognitive neuropsychology and philosophy of mind. British Journal for the Philosophy of Mind, 44, 589-622.
- Sabbagh, M. A., & Taylor, M. (2000). Neural correlates of theory-of-mind reasoning: An event-related potential study. Psychological Science, 11(1), 46-50.
- Strawson, P. F. (1959). *Individuals: An essay in descriptive metaphysics*. New York: Routledge.