Available online at www.sciencedirect.com
SCIENCE DIRECT
Brain and Cognition 60 (2006) 272–281
Brain and Cognition
www.elsevier.com/locate/b&c

Cognitive neuroimaging: Cognitive science out of the armchair

Greig I. de Zubicaray *

Centre for Magnetic Resonance, The University of Queensland, Qld, Australia

Accepted 3 June 2005
Available online 9 January 2006

Abstract

Cognitive scientists were not quick to embrace the functional neuroimaging technologies that emerged during the late 20th century. In this new century, cognitive scientists continue to question, not unreasonably, the relevance of functional neuroimaging investigations that fail to address questions of interest to cognitive science. However, some *ultra-cognitive* scientists assert that these experiments can never be of relevance to the study of cognition. Their reasoning reflects an adherence to a functionalist philosophy that arbitrarily and purposefully distinguishes mental information-processing systems from brain or brain-like operations. This article addresses whether data from properly conducted functional neuroimaging studies can inform and subsequently constrain the assumptions of theoretical cognitive models. The article commences with a focus upon the functionalist philosophy espoused by the ultra-cognitive scientists, contrasting it with the materialist philosophy that motivates both cognitive neuroimaging investigations and connectionist modelling of cognitive systems. Connectionism and cognitive neuroimaging share many features, including an emphasis on unified cognitive and neural models of systems that combine localist and distributed representations. The utility of designing cognitive neuroimaging studies to test (primarily) connectionist models of cognitive phenomena is illustrated using data from functional magnetic resonance imaging (fMRI) investigations of language production and episodic memory.

Keywords: Functional magnetic resonance imaging; Connectionism; Cognitive science; Language production; Episodic memory; Neuroimaging

1. Introduction

Consider the following statements made recently by Coltheart (2004a) in relation to functional neuroimaging: "…facts about the brain do not constrain the possible natures of mental information-processing systems. No amount of knowledge about the hardware of a computer will tell you anything about the nature of the software that the computer runs. In the same way, no facts about the activity of the brain could be used to confirm or refute some information-processing model of cognition" (p. 22). This position, espoused by *ultra-cognitive* psychologists (so-termed by Shallice, 1988), is an expression of the functionalist philosophy articulated prominently by Fodor (1981) and Pylyshyn (1984). Functionalism assumes information processing occurs at a level of abstraction that does not depend on the physical composition of the system. Abstract information processing depends only upon the *organization* of the mental system—the relationships among parts that are definable according to the function they perform—and these are presumed to operate according to discernible psychological rules or principles.

The adoption of the software analogy for the mind occurred as cognitive science developed an interest in computational modelling (e.g., Block, 1980; Pylyshyn, 1984). Block (1980) referred to the resulting conflation of functional and computational descriptions as *computation-representation functionalism.* In an influential publication on visual perception, Marr (1982) proposed three different levels for machine implementations of information processing; computational theory, representation and algorithm (i.e., software), and hardware implementation. He also emphasised the independence of the former two levels from the latter, while acknowledging their

* Fax: +617 3365 3833.
E-mail address: greig.dezubicaray@cmr.uq.edu.au.

0278-2626/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandc.2005.11.008

interconnections. Later, Block (1995) would write a book chapter with the unambiguous title "The mind as the software of the brain," reinforcing the primary research focus of modern cognitive science.

It is worth noting that the functionalist perspective reflects only one attempt at a solution to the perennial mind-body problem in philosophy. Despite the occasional antipathy expressed among philosophers, such solutions are not dogma, and they should not be revered as such. Many scientists, including cognitive scientists, adhere to materialist philosophies that link cognitive and brain processes. Some, such as Dennett (1991a, 1991b), provide strong countering arguments to functionalism. Like Dennett (1991a), I think there is something ludicrous about a position that precludes scientists from attempting empirical explorations of models because others claim to have an a priori proof that all such attempts are hopeless. By precluding functional neuroimaging investigations from providing evidence for or against cognitive models, ultra-cognitive science is adopting the position of armchair philosopher: ready to observe, ready to berate, never ready to engage.

Functional neuroimaging is not the only experimental method to be deemed irrelevant by the ultra-cognitivists. Over two decades ago the connectionist movement showed promise to move cognitive modelling closer to neural modelling. This is because the constituents of its models are nodes in networks connected in ways that resemble brain networks (Dennett, 1991b; Medler, 1998). Seidenberg (1993) noted connectionism's potential to extend modelling from being merely descriptive to the level of providing explanatory theories. However, as recently as last year, Harley (2004a) had cause to lament that connectionism appeared largely ignored by cognitive scientists.

Like functional neuroimaging, connectionism eschews an arbitrary distinction between software and hardware, thereby conflating Marr's (1982) proposed levels of machine implementation. Within connectionist models, hardware and software are inexorably intertwined (nodes are neuron-like), networks develop pragmatically, representations are discovered, *learning occurs*. No assumptions are made regarding governing psychological rules or principles in the functionalist manner, rather connectionist models demonstrate how the cognitive systems become organised the way they are (see Bechtel & Abrahamsen, 2002; Medler, 1998).

That connectionist modelling and neuroimaging technologies emerged at roughly the same time in history is an interesting coincidence. That both methods of empirical exploration with their similar emphases could be largely ignored or labelled irrelevant by a group of ultra-cognitive scientists is not coincidental: They represent alternative approaches to a functionalist-inspired cognitive science. More specifically, they represent attempts to apply *unified models* of cognitive and neural processes. As with any type of model, it is intended that they be tested.

1.1. Localist representations and the straw man

Cognitive psychologists continue to debate the value of including localist representations in their models. Despite the debate, it seems that non-connectionist modellers tend to adopt localist representations more frequently than their connectionist counterparts (primarily because they tend to eschew distributed processing in their models). I use the term "non-connectionist" here in the same manner as Coltheart (2004a), referring to largely descriptive computational models that do not a priori describe their connections as being brain- or neuron-like (apparently to preserve the functionalist's theoretical distinction between mental and physical systems—see the section on models of language production below).

For example, Coltheart (2004b) recently offered evidence in support of models that comprise systems of localist representations of word (phonological and orthographic) and object (visual/structural) forms, i.e., *lexicons*. He argued that models comprising solely distributed representations fail to account for the range of behavioural observations from lesion patients. At its outset, connectionism made the central claim that knowledge is coded in a distributed manner (Rumelhart, McClelland, & the PDP Research Group, 1986). However, localist representations *do* emerge in connectionist models, and more recent approaches to connectionist modelling have explicitly incorporated localist representations (see Bowers, 2002; Page, 2000). In fact, preempting Coltheart (2004b), Bowers (2002) showed the deficiencies of purely distributed models of language, and the advantages of incorporating localist representations within a connectionist architecture. A considerable number of connectionist models now employ a combination of localist and distributed representations (see Page, 2000).

How is this debate about localist and distributed representations relevant to the debate about functional neuroimaging? The activation patterns of nodes in connectionist models can be recorded to determine whether localist representations have emerged during processing, analogous to neuroimaging voxel timeseries data (e.g., Medler, Dawson, & Kingstone, 2005). These nodes can be thought of as consisting of a single neuron or a distinct population of neurons (Page, 2000). While the complexity of these models cannot approach the complexity of brain networks, the approach is nevertheless similar. Modern functional neuroimaging methods, such as functional magnetic resonance imaging (fMRI; with which I am most familiar), provide neural signal timeseries information during the performance of tasks designed to engage a cognitive process of interest. In 1997, Marcus Raichle suggested that the construction of cognitive paradigms is the "real Achilles' heel" in functional neuroimaging experiments. This is still the case. However, contemporary experiments frequently use identical experimental designs to those enjoying the endorsement of cognitive science, with some obvious constraints introduced by the scanner environment. The neuroimaging timeseries data are treated simply as another