ON FUNCTIONS AND MECHANISMS IN THE INVESTIGATION OF COGNITIVE CAPACITIES

One influential tradition of thought in philosophy of psychology holds that explanations of cognitive capacities such as working memory, stereopsis, and emotion proceed by showing that these complex capacities are made up of simpler sub-capacities organized together so that they exhibit the explanandum phenomenon. This sort of explanation is usually referred to as functional analysis. In contrast, contemporary views about the nature of neuroscientific explanation maintain that good neuroscientific explanations describe mechanisms which reveal the causal structure of the world. In what follows I argue that functional analysis is a form of mechanistic explanation. More precisely, I take it that functional analysis is a mechanism sketch which omits various details about the mechanisms under study, but which turns into a complete mechanistic explanation once these details are appropriately filled in. I argue against the received view about the relationship between psychological and neuroscientific explanations by showing that neither the distinctness, nor the autonomy thesis can face the challenges raised against them. I conclude that while both functional analyses and neuroscientific mechanisms are explanatory relevant, the former are best understood as elliptical mechanistic explanations. The proposed solution suggests a framework for integrating psychological and neuroscientific accounts of cognitive capacities.

The present argument is organized in the following manner. I start by introducing the causal-mechanical account of neuroscientific explanation defended by Craver (2007) and claim that it constitutes a promising alternative to the more traditional reductive view of mechanistic explanation (Churchland 1989; Bechtel and Richardson 1993). Since mechanistic explanation is a form of constitutive explanation, I discuss the two most influential traditions which account for this type of explanation: the reductive (Churchland 1986; Kim 1992) and the systems tradition (Fodor 1974; Cummins 1983). On the system tradition, psychological capacities are to be explained functionally, that is by what is often called functional analysis. The received view is that functional analysis is autonomous and thus distinct from mechanistic explanation (Fodor 1968; Cummins 1983; 2000). But if psychological explanation is functional and neuroscientific explanation is mechanistic (Craver 2007), and moreover functional analysis is autonomous and distinct from mechanistic explanation, then psychological explanation is distinct and autonomous from neuroscientific explanation. The main purpose of the present argument is to show that functional and mechanistic explanations

are not distinct and autonomous precisely because functional analysis is a form of mechanistic explanation, namely an elliptical mechanistic explanation - i.e., a mechanism sketch.

Mechanistic explanation is a form of constitutive explanation which describes the capacities (functions, behaviours, activities) of a system as a whole in terms of its components, their properties and the way they are organized together (Bechtel and Richardson 1993; Machamer, Darden et al. 2000). Therefore, constructing a mechanistic explanation requires decomposing the capacities of the whole mechanism into subcapacities. This characterization of the mechanistic model of explanation places it in the continuation of the systems tradition which holds that explanation of a cognitive capacity proceeds by functional analysis. However, in contrast with functional analysis, a mechanistic explanation assigns the purported capacities to so-called structural components of a system (as opposed to its functional components). Thus, despite their similarities, psychological explanations that appeal to functional decompositions and neuroscientific explanations that describe mechanisms have traditionally been considered to be distinct and autonomous from one another.

I claim that neither the autonomy, nor the distinctness theses can be reasonably defended once one takes into account the robust notion of mechanism put forward by contemporary views about the nature of neuroscientific explanation. By arguing that the main kinds of functional analysis (i.e., task analysis, functional analysis by internal states, and boxology) can be interpreted as elliptical mechanistic descriptions of particular psychological capacities, I seek to show that the dichotomy between reductionism and autonomy of psychological explanations from neuroscientific explanations is false. I take it that the idea of redefining functional analysis as a form of mechanistic explanation suggests an alternative way of integrating psychological and neuroscientific explanations in a unified science of cognition.

Although it would be false to state that all explanations describe mechanisms, most of them do. This is why I took the present paper to explore the possibility that certain psychological explanations which are traditionally held to be purely functional might be a step in a mechanistic explanation. I think that the proposed arguments suggest that to describe an item/capacity functionally is, ipso facto, to describe its contribution to a mechanism. Functional analyses are elliptical mechanistic explanations which appropriately constrained might entail mechanistic schemata and ideally complete mechanistic explanations which would describe both the functional and the structural properties of a given mechanism. The

integrationist view entailed by these conclusions is not entirely new. Many psychologists have been moving away from less mechanistically constrained models towards models that take more and more neuroscientific constraints into account (e.g., Kosslyn, Thompson et al. 2006; Gazzaniga 2009).

Recognizing that constitutive explanation does not split in two distinct forms of explanation, i.e., functional analysis and mechanistic explanation, is also a first step in finding the right tools to evaluate hypotheses put forward about the nature of our psychological capacities. I take it that if the present argument is on the right track, it might point towards new resources for dealing with particular problems raised both in philosophy of mind and philosophy of science, such as questions about top-down and bottom-up causation, natural kinds and eventually the representational thesis.

References:

Bechtel, W. and R. C. Richardson (1993). Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research. Princeton, N.J.: Princeton University Press.

Churchland, P. M. (1989). A Neurocomputational Perspective: The Nature of Mind and the Structure of Science, Cambridge, Mass.: MIT Press.

Churchland, P. S. (1986). Neurophilosophy: Toward A Unified Science of the Mind-Brain, Cambridge, Mass.: MIT Press.

Craver, C. F. (2007). Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience, Oxford: Oxford University Press.

Cummins, R. (1983). The Nature of Psychological Explanation. Cambridge, Mass.:MIT Press.

Cummins, R. C. (2000). '"How Does It Work" versus "What Are the Laws?": Two Conceptions of Psychological Explanation' *Explanation and Cognition*, 117-145. F. Keil and R. A. Wilson, Cambdrige, Mass.: MIT Press.

Fodor, J. A. (1968). 'The Appeal to Tacit Knowledge in Psychological Explanation', *Journal of Philosophy*, **65**(October): 627-640.

Fodor, J. A. (1974). 'Special sciences', *Synthese*, **28**(2): 97-115.

Gazzaniga, M. S. (2009). The Cognitive Neurosciences, Cambridge, Mass.: MIT Press.

Kim, J. (1992). 'Multiple Realization and the Metaphysics of Reduction', *Philosophy and Phenomenological Research*, **52**(1): 1-26.

Kosslyn, S. M., W. L. Thompson, et al. (2006). The Case for Mental Imagery. New York: Oxford University Press.

Machamer, P. K., L. Darden, et al. (2000). 'Thinking about Mechanisms', *Philosophy of Science* **67**(1): 1-25.