MECHANISMS, DECOMPOSITION AND THE SOCIAL SCIENCES

Section: Philosophy of Science

SHORT ABSTRACT

According to mainstream literature in philosophy of science, mechanistic explanations are essentially machine-like explanations – they assume a composite set of interrelated parts, each one of them performing its own function and producing the overt behaviour of the system (Bechtel & Richardson, 1993 [2010]). The research strategy of decomposition and localization is crucial to fulfil the mechanistic project. I argue however that this framework is ill suited to accommodate the role of mechanisms in the social sciences. I claim that mechanistic explanations in the social sciences search for identifying recurrent patterns of interaction rather than penetrating the inner structure of phenomena. I use some examples of game theoretic modelling to illustrate this point.

LONG ABSTRACT

According to mainstream literature in philosophy of science, mechanistic explanations are essentially machine-like explanations – they assume a composite set of interrelated parts, each one of them performing its own function and producing the overt behaviour of the system (Bechtel & Richardson, 1993 [2010]). There are different ways in which these components interact. An important aspect to take into account is the degree of decomposability (and *exchangeability*) of the components. Suppose we want to explain the functioning of a car. We can decompose it in different parts, each one of them dedicated to carrying out a specific task. The engine, the fires, and the wheels are discrete components, and each one of them contributes to the functioning of the car in a specific way. The essential feature of this system is that the components are not fungible because they are task-specific – each one of these components is assigned a fixed, discrete role within the whole system, and it cannot be used for a different purpose. The fires cannot be used, for instance, to produce combustion.

The research strategy of decomposition and localization is crucial to fulfil the mechanistic project in decomposable systems (Bechtel & Richardson, 1993 [2010]; Bechtel, 2006 Craver, 2007; Wimsatt, 2007). Decomposing is tantamount to identifying the sub-task of the system, and localization consists precisely in identifying the material that implements such function. Once a given function is identified as realized or performed by a given organ, part or component of the system, we should continue more deeply searching for the still deeper sub-functions and sub-components that are fulfilling the aforementioned task. Decomposition and localization is a process of continuous search for more detailed and deeper micro-components. Note that this research strategy makes sense only in relation to a specific object or phenomena; the mechanism implementing cell respiration, and the inner structure identified to accomplish it, cannot be extrapolated to account for digestion. Each parcel of reality has an inner structure that must be described and that cannot be transposed.

Examples of decomposable models abound in science, especially in biology, chemistry, physics and psychology. I argue however that the decomposition and localization strategy is ill suited to account for the role of mechanistic explanations in the social sciences. In the social sciences non-decomposability is rift. The high-level behaviour of social systems is the result of the interaction of its components, but these components are not responsible for performing discrete, non-transferable sub-tasks. The limiting case takes place when all the sub-components are simply exchangeable: in a way all of them are doing the same thing. A good example is Schelling's model of segregation (1969). Schelling placed pennies and dimes on a chessboard and moved them around according to various rules. He interpreted the board as a city, and individuals of two different groups (black and whites, boys and girls, etc.) with dimes and pennies. He distributed these pennies and dimes randomly across

the board. Schelling later assumed that the pennies were "unhappy" if they had more than four dimes as neighbours, and the other way round, dimes were "unhappy" if they were surrounded by more than four pennies. Pennies having more than four dimes as neighbours, moved to a new location where there were fewer dimes, and the same occurred for dimes. The upshot of the process is the generation of segregation: dimes and pennies end up being concentrated in different parts of the chessboard. Note that the properties of the whole system are dependent on the behaviour of the components, but these components are homogeneously performing the same task – to wit, avoiding having more than four neighbours differing from themselves.

In non-decomposable systems the inner structure of the components is irrelevant because their behaviour is homogeneous. What is crucial is the relational interaction with the other components of the system. Note that the behaviour of the system as a whole depends on the behaviour of its parts, but the sub-task performed by the sub-components is almost alike; there is no specialized, dedicated sub-components fulfilling non-transferable tasks. Systems like these are highly homogeneous, and they do not display modularity. Non-decomposable systems have a flat ontology, and the research strategy of decomposition and localization is essentially irrelevant and fruitless for them.

Game theoretic models are crucial for illustrating this insight. In fact game theoretic models search for identifying social mechanisms. The goal is to identify recurrent, abstract patterns of interaction that can be abstracted away from the material basis implementing them. Within the context of game-theoretic models, uncovering the inner structure of phenomena is simply irrelevant. The objective of mechanistic explanations in the social sciences is to bridge the gap between different phenomena using - to this intent - a conceptual machinery that is highly abstract and idealized. Whereas the decomposition and localization strategy ultimately identified mechanisms that could not be transposed from one system to another, the goal of game-theoretic model is precisely the reverse – to identify abstract patterns of interactions that are independent from the objects implementing them. Thus these patterns of interaction can be applied to humans, to animals or even to bacteria or firms. Within this context, the strategy of decomposition and localization seriously misses the point – there are in fact no specific, object-dependent mechanisms to look for, but rather abstract mathematical structures that can be substantiated upon different phenomena.

References

Bechtel, W. (. (2006). *Discovering Cell Mechanisms: The Creation of Modern Cell Biology.* Cambridge, MA: Cambridge University Press.

Betchel, W., & Richardson, R. (1993). *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research.* Cambridge, M.A.: MIT Press.

Craver, C. (2007). *Explaining the Brain: Mechanisms and th Mosaic Unity of Neuroscience.* Oxford: Clarendon Press.

Fodor, J. (1983). *The Modularity of Mind.* Cambridge: MIT Press.

Giere, R. (1988). *Explaining Science: A Cognitive Approach.* Chicago: University of Chicago Press.

Schelling, T. (1969). Models of Segregation. *American Economic Review , 59 (2)*, 488-483.

Wimsatt, W. (2007). *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality.* Cambridge: Harvard University Press.