Calm down dear, it's only supervenience

Jason Noble, Eric Silverman, Manuel de Pinedo García

In the beginning everyone agreed that the business of the mind was conducted using propositional logic on internal representations. Then Rumelhart and McClelland (1986) resurrected connectionism, the idea that simple neuron-like units were a viable cognitive architecture. Enthusiasm for the new paradigm led to claims that the view of mind as a propositional-logic machine was outmoded. Jerry Fodor -- perhaps understandably so, having written a book entitled *The Language of Thought* -- was perturbed by the implication that such a language was a chimera that could be replaced by a sufficiently complex neural network.

Fodor and Pylyshyn (1988; hereafter F&P) therefore challenged the connectionists to show how their approach could possibly produce a systematic intelligence. The challenge was largely evaded rather than answered, in that no 1980s-connectionist architectures achieved systematicity. Subsequently, a "post-connectionist" movement emerged, combining enactivism, behaviour-based AI, embodied and distributed cognition, dynamical systems theory, etc. These ideas share anti-dualist and anti-representational themes (see Thompson 2010 and García Rodríguez & Calvo Garzón 2010 for recent examples).

Before deciding how the new connectionism stands up to the systematicity challenge, we must ask: what is the core of F&P's argument? We propose that their argument is centred around the supervenience relation. In other words, we can do science at L1, e.g., neuroscience, or at L2, e.g., representational cognitive psychology, or we can look at how L1 *produces* or *gives rise* to L2. The strong version of the F&P objection is "if you can't show me how L2 supervenes on L1, then your proposed L1 is illegitimate." But that doesn't make much sense: why should the duty to demonstrate a possibly mysterious supervenience relation fall solely on the L1 side? F&P are presumably happy with the idea that L2 supervenes on at least *something* physical. Thus it is better to read them as merely saying that the L1--L2 supervenience relation is non-obvious and worthy of further research. We can also take them to be defending against the strong opposing claim that L2 is unnecessary or reduces completely to L1.

We assume that explanatory pluralism is inevitable in the cognitive sciences and beyond; we take this to be uncontroversial. We have argued elsewhere (de Pinedo García & Noble, 2008) that in explaining the behaviour of evolved agents, both agent- and sub-agent-level explanations will be necessary. We are not saying that Fodorian representationalism is specifically the correct theory at L2, only that the presence of a suitably compelling L1 is not going to obviate the need for some variety of L2 explanation.

Does the success of an approach such as the new connectionism depend on it producing a bridging explanation between itself, L1, and some L2? Perhaps not. Consider an example from artificial life -- a particularly useful methodology for exploring potential bridging explanations

between two levels of description. Quinn (2002) took a step back from Shannon and Weaver's (1949) characterisation of communication as requiring a signaller, a receiver and an information channel. He constructed a model in which agents seeking to complete a task which would benefit from communication have no pre-existing signalling channel. In Quinn's model the agents evolved to convey information to one another through sensorimotor interaction, demonstrating how a communication channel could emerge.

At first sight this seems potentially earth-shaking: a model which demonstrates that communication does not require a signaller, a receiver and a channel looks quite damaging to Shannon's framework. However, the existence of the model does not mean that Shannon's framework is without value. Having seen that it is possible for signalling and response behaviours to emerge from sensorimotor interactions (L1), we may still find Shannon's picture of communication useful when discussing explanations of communication at a higher level (L2).

Essentially, Quinn has followed F&P's prescription for avoiding the explanatory fascism of the early connectionists: he has provided a new insight into how L1 becomes L2, but not eliminated the need for a discussion of L2 itself. The lesson from the Quinn example is that the new connectionism, just like the old connectionism, should not be expected to explain everything. It is a very reasonable scientific question to ask how we get from L1 (neurons, feedback cycles, dynamical systems, etc.) to L2 (concepts, reasoning, propositional logic, or at least the appearance of such things in human cognition). It is also reasonable to conduct research wholly located at either level. Neither side, in an L1 versus L2 debate, can use the absence of a bridging explanation as a stick to beat the other with.

To the extent that a researcher from the new connectionist paradigm claimed that their L1 description (e.g., a dynamical system) would obviate the need for L2-talk (e.g., representations), we would turn the question back on them and ask how they account for the pragmatic usefulness of alternative levels of description such as the folk-psychological, the Fodorian view of the mind as a propositional logic machine, and the Rylean view of agents working in a domain of reasons. The primary error of the 1980s connectionists was that their understandable enthusiasm for their new paradigm led them to make category errors in what they claimed it could explain. It would be a shame for the new connectionists to go down the same path.

The question of how something like representations may emerge from a neural substrate is a tough problem, but it is a *scientific* problem. To our knowledge, the best guess at the chain of developments necessary to go from the simplest cognition to complex human behaviour would be Braitenberg's (1984) "Vehicles", but that is of course a speculative story. Our point is that we do not have to answer the question here, and that such questions are better suited to the laboratory rather than the armchair. Empirical investigation may see some levels of description fall away and new ones emerge: this is progress. We would echo Quine (1951) in noting that our choices of explanatory levels and associated supervenience relations are "where rational, pragmatic."

Braitenberg, V. (1984). *Vehicles: Experiments in Synthetic Psychology*. MIT Press.

de Pinedo García, M. and Noble, J. (2008). Beyond persons: extending the personal / subpersonal distinction to non-rational animals and artificial agents. *Biology and Philosophy, 23*(1), 87-100.

Fodor, J. A. (1975). *The Language Of Thought*. Crowell Press.

Fodor, J. A., and Pylyshyn, Z. (1988). Connectionism and cognitive architecture: a critical analysis. *Cognition*, 28: 3-71.

García Rodríguez, Á. and Calvo Garzón, P. (2010). Is cognition a matter of representations? Emulation, teleology, and time-keeping in biological systems. *Adaptive Behavior*, 18(5):400-415.

Quine, W. V. O. (1951). Two dogmas of empiricism. *Philosophical Review*, 60(1):20-43.

Quinn, M. (2001). Evolving communication without dedicated communication channels. In Kelemen, J. and Sosik, P., editors, *Advances in Artificial Life: Sixth European Conference on Artificial Life (ECAL'01)*, vol. 2159, 357-366. Springer.

Rumelhart, D. E., McClelland, J. L., and the PDP Research Group. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, volume 1. MIT Press.

Shannon, C. E. and Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press.

Thompson, E. (2010). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Belknap Press.