

Final draft of paper published in *International Journal of Philosophical Studies*, 16(3), 371-6, 2008.

## **Minimal Representing: a Response to Gallagher**

**Michael Wheeler**  
**Department of Philosophy**  
**University of Stirling**

In his contribution to this volume, Shaun Gallagher casts a sharply focussed critical eye over positions which claim that action is to be explained, in part, by appeal to minimal representations. One of the positions to come under fire in this way is a view that I developed in my book *Reconstructing the Cognitive World* (Wheeler, 2005). In this brief response I shall try to defend my own brand of representational minimalism against some of Gallagher's worries.

Following an analysis of the minimalist's commitments, Gallagher offers a six-point negative characterization of minimal representation, that is, six things that minimal representations are (apparently) *not*. They are not (1) wholly internal, (2) discrete, identifiable, enduring things, (3) passive, (4) decoupleable, (5) strongly instructional, or (6) homuncular. The minimalist's aim is to pare down the concept of representation so as to render it suitable for the explanation of real-time action. Gallagher's analysis suggests that, in pursuing this goal, the minimalist divests the concept of representation of properties 1-6. But, argues Gallagher, properties 1-6 are the very properties that, in some combination, make a concept a genuinely representational one. So, because minimal representations possess none of these properties, "the idea of a minimal representation no longer conforms to the criteria that would make it a representation" (Gallagher, this volume). In the end, then, minimal representations are not representations at all.

Let's call properties 1-6 *Gallagher-properties*. One way for the minimalist to respond to the challenge here would be to show that the states or processes that are candidates for minimal representation-hood do, pace Gallagher, possess one or more (the more the better) Gallagher-properties. In my way of thinking, minimal representation is action-oriented representation, where an action-oriented representation is one that is (i) *action-specific* (tailored to a particular behaviour and designed to represent the world in terms of specifications for possible actions), (ii) *egocentric* (features bearer-relative content as epitomized by spatial maps in an egocentric co-ordinate system), and (iii) *intrinsically context-dependent* (the explicit representation of context is eschewed in favour of situated

special-purpose adaptive couplings that implicitly define the context of activity in their basic operating principles). (For a worked-through example of action-oriented representation that illustrates these features, see the discussion of Franceschini et al.'s (1992) 'fly robot' in my paper in this volume.) Given this characterization of minimal representation, I agree without qualification that such elements do not possess Gallagher-properties 3 and 5. Moreover, since the relationship between decoupleability (Gallagher-property 4) and representation seems to get murkier every time I think about it, and since I am at least convinced that decoupleability is not necessary for minimal representation (see Wheeler, 2005), I propose to ignore that issue for today. (The properties of decoupleability and being strongly instructional will, however, crop up again, in relation to different points.) That leaves three Gallagher-properties to be discussed.

As Gallagher observes, my account of representation is based, in part, on what I call the *neural assumption*, which states that if intelligent action is to be explained in representational terms, then whatever criteria are proposed as sufficient conditions for representation-hood, they should not be satisfied by any extra-neural elements for which it would be unreasonable, extravagant, or explanatorily inefficacious to claim that the contribution to intelligent action made by those elements is representational in character. The justification for the neural assumption is largely methodological. It seems likely that neural states and processes do something that is, for the most part, psychologically distinctive, and we expect the concept of representation to help us explain how that something comes about. Thus there is a clear sense in which action-oriented representation is overwhelmingly brain-bound.

It does not follow from the neural assumption that, despite what I said earlier, minimal representation is, in some covert way, strongly instructional in character, that is, fully determinative of the behavioural outcome to which it contributes (Gallagher-property 5). The neural assumption as stated allows that non-representational causal contributions by the non-neural body and the external environment, to the fine-grained structure of intelligent action, may go beyond that of mere background conditions for processes of internal, representation-driven control. This point also clarifies the sense in which "Wheeler... [suggests]... that minimal representations involve aspects of a system that is brain, body... but also environment" (Gallagher, this volume). According to me, minimal representations *involve* factors in the non-neural body and the environment in the sense that the behavioural outcomes that the neurally located representational elements support will depend also on non-trivial causal

contributions from extra-neural factors. In most cases, however, those extra-neural factors will not themselves qualify as the vehicles of the representation itself. Another conclusion that does *not* follow from the neural assumption is that the cognitive is restricted to the neural. For one thing the neural assumption is liberal enough to allow *some* external factors (e.g. road-signs, certain linguistic structures) to qualify as representations in the sense relevant for cognitive-scientific explanation. For another, unless one identifies the cognitive with the representational, a move which, as Gallagher's discussion elegantly demonstrates, would be a mistake, then the neural may do something that is, for the most part, distinctive and distinctively representational, within the class of contributions that count as cognitive. So, for all the neural assumption says, the extended mind hypothesis (Clark and Chalmers, 1998) might still be true.

What does all this tell us? Gallagher claims that the fans of minimal representation want to do without Gallagher-property 1, that is, full internality. Evidence for this claim may be found in the work of, for example, Rowlands (2006) who holds that the vehicles of representation extend into the environment. However, the sense in which I agree that minimal representations are not fully internal is *heavily* qualified. For my endorsement of the neural assumption means that I do expect the overwhelming majority of minimal representations to be spatially located *inside* the agent – indeed, inside the brain. The extra-neural factors that qualify as minimal representations are essentially friendly interlopers into what is largely an inner sanctum. And since I think that if one's theory of representation permits regular violations of the neural assumption, one simply doesn't have a good enough theory of representation, I treat compliance with the neural assumption, and thus with a qualified internality constraint, as precisely the kind of criterion for representation-hood that Gallagher's negative characterization suggests that I should not. This enables me to place some distance between my own view and the claim that minimal representations do not possess Gallagher-property 1.

Do minimal representations possess Gallagher-property 2, that of being discrete, identifiable, enduring things? Recall that, for me, minimal representations are context-dependent egocentric control structures for situation-specific actions. I see no reason to think that such structures couldn't be discrete or identifiable, but could they be enduring? It is true that some flagship examples of minimal representations (e.g. the 'built-on-the-fly' snap maps discussed in my paper in this volume) enjoy only a transient existence as temporary structures built in the heat of real-time action. But others, it seems, may be stored structures that endure over time. For example, Mataric's sonar-driven mobile robot Toto

(Mataric, 1991), mentioned by Gallagher, deploys a navigational system that uses minimal (action-oriented) representations of spatial landmarks. The landmarks in question are encoded in terms of patterns of sensorimotor activity. Thus, if Toto keeps detecting proximally located objects on its right hand side, while its compass bearing remains unchanged, then a 'right-wall' is internally encoded, not as some objective entity, but in terms of the robot's sensorimotor 'experience' at the time. Histories of these structured sensorimotor 'experiences' are stored as connected nodes in a distributed graph. Later, using this graph, paths through the environment may be encoded as navigation-supporting sequences of past, current, and expected 'experiences'. Thus, whatever I may have inadvertently suggested in the past by concentrating my discussion on minimal representations whose existence is short-lived, there is reason to reject the claim that minimal representations could not be enduring things and so fail to possess Gallagher-property 2. Some minimal representations are transient structures – that's surely right – but some, it seems, endure.

Finally let's consider Gallagher-property 6, homuncularity. In my view systemic homuncularity is necessary for subagential representation, so minimal representations will certainly be associated with that property (Wheeler, 2005). A system is homuncular when (a) it can be compartmentalized into a set of hierarchically organized communicating modules, and (b) each of those modules performs a well-defined sub-task that contributes towards the collective achievement of the overall adaptive solution. Gallagher's first critical strategy here is to suggest that the very notion of homuncularity is conceptually problematic. He writes: homuncularity "seems to be something similar to Clark and Grush's emulator [see Gallagher's paper for a description of this model] *sans* decoupleability. At the same time it is not clear what off-line but not decoupled means..." (Gallagher, this volume). In order to see what is going on here, we need to understand how the concept of decoupleability is related to the on-line-off-line distinction.

An element or mechanism realizes the property of decoupleability just when the system in which it figures has been designed precisely so that that element or mechanism may become activated in the absence of its usual eliciting stimuli. In general terms, on-line intelligence is characterized by the production of fluid and flexible real-time adaptive responses to incoming sensory stimuli (e.g. playing squash, avoiding a predator). By contrast, off-line intelligence is characterized by thinking that is disengaged from the ongoing flow of perception and action (e.g. mentally planning one's squash match strategies, reflecting on the quality of the beer in Munich while sitting on a bus in Edinburgh). Given these definitions

Gallagher is right, of course, that decoupleability is necessary for off-line processing, and that explains why “it is not clear what off-line but not decoupled means”. Nevertheless, none of this seems to present any sort of problem for homuncular analysis. So why does Gallagher think otherwise? He seems to assume that if a system is homuncular then it involves off-line processing. If that were true then of course there would be a problem in cases where putatively homuncular systems do not realize the property of decoupleability. But why think that homuncularity results in cognitive processing that is *necessarily* off-line? Indeed, recall the previously mentioned fly robot. There is no sense in which the homuncular states and processes that underlie the generation of real-time adaptive navigation by this robot are disengaged from the ongoing flow of perception and action. Indeed, they are intimately embedded in that flow. In short, they are *on-line*. But now since the homuncular states and processes in question are on-line, there is no tension with the idea that they do not realize decoupleability. In sum, the notion of homuncularity appears to be in good conceptual order.

Gallagher proceeds to suggest that even if we can make conceptual sense of homuncular systems, and thus potentially of minimal representations, we will not find them in the mechanisms underlying real-time embodied action. Why? Because such mechanisms realize the “dynamic systems concept of a self-organizing continuous reciprocal causation” (Gallagher, this volume). Continuous reciprocal causation (as characterized by Clark, 1997) is causation that involves multiple simultaneous interactions and complex dynamic feedback loops, such that (a) the causal contribution of each systemic component partially determines, and is partially determined by, the causal contributions of large numbers of other systemic components, and, moreover, (b) those contributions may change radically over time. As I have argued previously (Wheeler, 2005), continuous reciprocal causation (CRC) undermines representational explanation because it undermines homuncularity; and it undermines homuncularity because it undermines modularity. Modularity is necessary for homuncularity, and thus for representation. CRC undermines modularity because, as CRC increases, it becomes progressively more difficult to specify distinct and robust causal-functional roles played by reliably reidentifiable parts of the system. The performance of any particular sub-task will increasingly be underpinned by larger and larger numbers of interacting components whose contributions are changing in highly context sensitive ways.

Gallagher and I agree, then, that real-time action may be underpinned by CRC, and that where this is the case there will be no place for minimal representations.

Where Gallagher and I disagree is that I see no compelling reason to follow him in thinking that *all* real-time action will be underpinned by CRC. For although this phenomenon looks to bestow a certain kind of useful large-scale adaptive flexibility on a system, it is rather less obvious that it introduces a set of properties that make mechanistically intelligible everything that we want to explain about real-time embodied action. For example, it is far from obvious (to me anyway) that an appeal to CRC alone has the resources to account for the core phenomenon of adaptive sensitivity to what is relevant *within* a context of action. A more compelling picture, I think, is the one sketched in my paper in this volume (see also Wheeler, 2005), a picture according to which CRC mediates the transitions *between* the sorts of situated special-purpose adaptive couplings that individually feature the property of intrinsic context-dependence. Once again our old friend the fly robot indicates that some of those couplings will be organized into homuncular systems and will feature action-oriented representations. I conclude, then, that minimal representations will be associated with Gallagher-property 6, homuncularity.

Gallagher's timely paper challenges the fan of minimal representation to be clear about what of a genuinely representational character remains, once the concept of representation has been stripped down far enough to enable it to play a useful and proper role in the explanation of real-time embodied action. This is without doubt a serious challenge, but it is one that, I think, can be met.

### **Acknowledgment**

This paper was prepared thanks to support granted by the AHRC, as part of project AH/F002963/1.

### **References**

- Clark, A. (1997) *Being There: Putting Brain, Body, And World Together Again*, Cambridge, Mass.: MIT Press.
- Clark, A. & Chalmers, D. (1998) 'The Extended Mind', in *Analysis* 58 (1), 7-19.
- Franceschini, N., Pichon, J. M., and Blanes, C. (1992) 'From Insect Vision to Robot Vision, *Philosophical Transactions of the Royal Society, series B*, 337: 283-94.
- Mataric, M. (1991) 'Navigating with a Rat Brain: a Neurobiologically Inspired Model for Robot Spatial Representation', in J.-A. Meyer and S. Wilson (eds.)

*From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior*, Cambridge, Mass.: MIT Press, 169-75.

Rowlands, M. (2006) *Body Language: Representation in Action*, Cambridge, Mass.: MIT Press.

Wheeler, M. (2005) *Reconstructing The Cognitive World: The Next Step*, Cambridge, Mass.: MIT Press.