

COMPARATIVE COGNITION & BEHAVIOR REVIEWS

Yesterday the Earwig, Today Man, Tomorrow the Earwig?

Christopher L. Hewitson, David M. Kaplan, and John Sutton

*Department of Cognitive Science, ARC Centre
of Excellence in Cognition and its Disorders,
and Perception in Action Research Centre
Macquarie University*

In this commentary, we highlight some relevant history of the situated cognition movement and then identify several issues with which we think further progress can be made. In particular, we address and clarify the relationship between situated cognition and antirepresentational approaches. We then highlight the heterogeneous nature of the concept of morphological computation by describing a less common way the term is used in robotics. Finally, we discuss some residual concerns about the mutual manipulability criterion and propose a potential solution.

*Keywords: robotics, morphological computation, antirepresentationalism,
mutual manipulability, causal specificity*

A Brief History of Situated Cognition Research

We applaud Cheng's effort to bring concepts from the framework of situated cognition to a wider biological audience, including researchers working on insect behavior. It is also delightfully ironic. Cheng notes that although the idea of "situated cognition has been bantered in philosopher and cognitive science for some time now . . . its connection with nonhuman animals has a more recent history" (p. 11). Although he is correct to point out that recent philosophical discussions have focused on the human case, much of the early history had a different focus, closer to home for the target audience of this article. For deep theoretical reasons, many of the earliest discussions of embodied cognition among researchers in artificial intelligence (AI) and robotics centered on the rich, adaptive behavioral repertoires of simpler organisms including insects. Therefore, in many ways, the discussion of situated cognition has truly come

full circle—from insects to humans and now back to insects again.

Consider, for example, roboticist Rodney Brooks's (1991) seminal article "Intelligence Without Representation," which perhaps more than any other single article served to launch the embodied cognition movement. A central goal of this article and much of Brooks's career was to build a research program in AI that drew inspiration from the intelligent behavior manifested by insects and other simpler biological organisms rather than humans. Brooks forcefully argued that researchers in AI had mistakenly assumed for too long that the hallmark of intelligence was disembodied reasoning or computation performed over explicit, language-like representations, instead of adaptive and flexible control of bodily action. According to Brooks, the latter is the real locus of natural (and so artificial) intelligence. This shift in

perspective, which Brooks (1999) later termed *behavior-based robotics*, led him and others to extract design principles for robots from a close inspection of the behavioral competencies exhibited by nonhuman organisms like ants and bees. The Brooksian perspective should resonate with biologists who naturally (and correctly) see intelligence and cognition through an evolutionary lens.

Brooks was by no means alone in embracing this perspective in the early development of the embodied cognition movement. In a seminal book that formed the foundation for much current work on embodied and extended cognition, Andy Clark (1997) used case studies of organisms like the “humble” cockroach to argue for an embodied view of intelligence. According to Clark (1997), the essence of intelligence and cognition is rooted in an organism’s basic capacities to “sense, act, and survive” (p. 4), and so can be found throughout the biological world. Barbara Webb’s (1994) influential robotic modeling of the mechanisms of cricket phonotaxis similarly emphasized how bodily structures can in some instances obviate the need for internal representation and computation in the service of intelligent behavior. Cognitive scientists in the 1990s thus had to take seriously the idea that the adaptive behavior exhibited by relatively simple mobile robots and insects should form the basis for scaled-up approaches to flexible decision making and action control in humans. The title of David Kirsh’s (1991) now famous critical response to Brooks (1991) succinctly (and somewhat sarcastically) captures the view: “Today the Earwig, Tomorrow Man?”

It is both interesting and exciting that debates about situated cognition (which, according to Cheng, subsumes embodied and extended cognition) have now come full circle. Early discussions revolved around insect models; “intermediate” discussions around humans; and at least some future discussions, one hopes, will revolve around insects again. Situated cognition research is in an important sense returning to its roots. With a nod to Kirsh (1991), we may summarize the past, present, and future of situated cognition as follows: Yesterday the earwig, today man, tomorrow the earwig.

Situated Cognition and Representationalism: Enmity, Friendship, or Neutrality?

Theorists in some areas of situated cognition, as Cheng notes, have embraced the revolutionary rhetoric by which Brooks and others, for a time at least, hoped to convince cognitive scientists to scale up from insects to humans without appealing to computations over representations. From the start, however, many argued that creatures like us require more sophisticated control systems because we have rangier and more conflicting goals, and face more representation-hungry problems in more complex environments. For Kirsh (1991) and Clark (1997, 2005), revised and nonclassical notions of dynamic and action-oriented representation would remain in the situated cognition theorist’s toolkit. From one perspective, the subsequent history of situated cognition can look like an unproductive sequence of standoffs between various strands of radical antirepresentationalism (enactivism, dynamical systems theory, direct realism), and various defensive bands of moderates who remained friends of the representational theory of mind. But this story of entrenched enmity between situated cognition and representationalism is partial and misleading.

Cheng sometimes identifies situated cognition, and especially its more “liberal versions,” with the claim that cognition is “fundamentally different from the standard cognition of representations” (p. 2). On this view, it is a “conservative” move to argue that accounts of complex cognitive phenomena are unlikely to “escape references to representations” (p. 12). But these ways of drawing the lines unhelpfully collapse two distinct issues. Situated cognition, we suggest, is not fundamentally a thesis about representations at all. The right target, instead, has always been *internalism* or *individualism*, the view that cognitive processes occur entirely in the individual head, fundamentally distinct and separate from body and environment. Although issues about mental representation have often been confused with issues about individualism, they are orthogonal (Sutton, 2015; Wilson, 1994). Nothing about the truth or falsity of internalism is settled by adopting a particular view about mental representation.

By treating situated cognition as neutral with respect to representationalism, we open up appropriate space for empirical inquiry to address a range of possible relations between internal and external resources in different cognitive systems. Sometimes the relevant computations occur over *external* representations in public symbol systems. Intelligent behavior and cognition can reflect many different combinations of neural, bodily,

Author Note: David M. Kaplan, Department of Cognitive Science, 16 University Drive, Australian Hearing Hub, Level 3, Macquarie University, NSW 2109, Australia

Correspondence concerning this article should be addressed to David M. Kaplan at david.kaplan@mq.edu.au.

environmental, and social resources. Empirical study is needed to identify the different ways in which such heterogeneous but complementary resources are integrated across distinctive cognitive ecologies (Hutchins, 2010; Menary, 2007; Sutton, 2015). To the extent that the situated cognition movement is radical, it is not because it dispenses in any blanket fashion with computations and representations but because it does not assume a solely internal location for cognitive states and processes in the heads of individual organisms.

This slight corrective to Cheng's tendency to *identify* situated cognition with antirepresentationalism in fact highlights some of his most promising claims about the possibility of a "merger" between situated and representational explanations (p. 10). As he notes, it is far from obvious that nonrepresentational explanation will prove sufficient to explain even insect cognition in its more complex and demanding forms. But this does not mean that situated cognition fails in these cases. Body and environment may still be playing key roles in larger cognitive systems, even if the internal components of the processes in question are computational. Over the timescales of cultural evolution and development alike, it may be that those internal computational processes are deeply transformed and shaped by the situated nature of biological intelligence.

Morphological Computation: A Heterogeneous Concept

The concept of *morphological computation*, which Cheng introduces in his interesting discussion of octopus fetching movements, might be one basis of such transformation. Morphological computation is the notion that certain processes can be carried out by the body that would otherwise be handled through computations performed by the central nervous system (CNS; Pfeifer & Bongard, 2006; Pfeifer, Iida, & Lungarella, 2014; Pfeifer, Lungarella, & Iida, 2007). As Cheng notes, claims about morphological computation provide especially clear examples of embodied cognition hypotheses. But the term is ambiguous, and several distinct phenomena are collapsed under a single concept (Hoffmann and Müller, 2017). In this section, we try to pull them apart.

Morphological computation most commonly describes a *reduction* in the computational load placed on the CNS by exploiting material properties of the organism's body such as its shape, structure, and dynamics. In this sense, bodily properties effectively change—and simplify—the computation to be performed. This

is what Cheng means when he uses the term, and this is why he prefers to call it "decomputation." Instead of trying to solve the computationally demanding inverse kinematics problem (mapping a desired outcome into motor commands; Flash & Sejnowski, 2001) for all the degrees of freedom of one of its hyperflexible arms, the octopus temporarily reconfigures its arm into a stiff quasi-jointed structure to transfer an object from one place to another (Sumbre, Fiorito, Flash, & Hochner, 2006). More specifically, precise patterns of muscle activations function to dynamically set joint locations and divide the arm into proximal, medial, and distal segments, which drastically reduces the degrees of freedom in the soft-body arm from near infinite to just 7. This vertebrate-like quasi-joint strategy in turn reduces the computational load on the associated neural circuitry and greatly simplifies the motor control problem to be solved. We therefore agree with Cheng that it is probably more appropriate to call this *decomputation* because bodily properties are harnessed to reduce the computational demands on the CNS rather than it being the case that computations are literally being performed in the non-neural body.

But Cheng's brief discussion of morphological computation, especially his decision to rename it, hints that there might be another—more literal—way to use the term. We agree. Robotics researchers do in fact seem to use the term to describe a *genuine division of computational labor* across the CNS and body, where the overall amount of computation to be performed remains essentially unchanged, but parts of the body literally do some of the computational processing. In this case, neural computation would quite literally be *offloaded* onto the non-neural body. Although Cheng provides some hints that there must be something more to the idea of morphological computation beyond that of decomputation, he offers no examples of *genuine* morphological computation. We briefly describe one kind of example proposed in the robotics literature to illustrate this alternative conception and merely raise the possibility that analogous biological examples may be found.

In a series of network simulations, Hauser and colleagues (Hauser, Ijspeert, Fuchslin, Pfeifer, & Maass, 2011, 2012) provided evidence that nonrigid or compliant physical bodies or body parts (modeled as recurrent networks of mass-spring systems) can be used to implement certain nonlinear computational operations. More specifically, they argued that the nonlinear input-output transformation achieved by these networks is implemented by the interconnected mass-springs that function

as the network weights. Critically, it is the morphological structure and dynamic properties of these mass-spring systems themselves that are supposed to provide the nonlinearity. Along similar lines, Nakajima et al. (2013) and Hoffmann and Müller (2017) maintained that the complex nonlinear dynamics of soft body structures such as the octopus arm might serve as computational *reservoirs*, high-dimensional dynamic systems, that can be exploited to perform nonlinear computations.

As Cheng highlights, morphological computation is a fruitful concept for researchers interested in situated cognition. But, as we have noted, it is also heterogeneous. We agree with Cheng that the quasi-articulated octopus arm is a case of morphological *decomputation* because the non-neural body reduces the computational load on the CNS but does not itself play a direct computational role. However, the theoretical work just cited suggests a stronger or more literal kind of morphological computation in which the non-neural body actually performs computations, although it remains to be seen whether this kind of computation occurs in real biological systems.

Mutual Manipulability: New Applications and Next Steps

It is encouraging to see that the mutual manipulability criterion (MM) is finding application in biology (Japyassú, 2017; Japyassú & Laland, 2017). The explicit goal in the original publication by Kaplan (2012) was to shift the direction of the debate about the embodiment and extension away from what he called “proprietary demarcation criteria” for determining the boundaries of cognition, which require problematic assumptions about the nature of cognition. MM was proposed as a way to steer the discussion down more fruitful and empirically grounded paths. Because MM reflects general interventional strategies used by scientists to experimentally test and determine mechanism boundaries, there was always latent potential for wider application of these ideas beyond the initial context of *human* cognition. Now that MM has proven useful for probing questions about extended cognition in spiders (Japyassú & Laland, 2017), it will be interesting to see how widely it can be applied to other biological taxa.

MM thus appears to be useful generally for detecting when some bodily or environmental feature should count as a real component as opposed to serving merely as a causally relevant background condition. But there is a residual worry that MM is still not restrictive enough,

if MM is satisfied merely, as Cheng puts it, “when causal influence flows both ways, from object to brain and from brain to object” (p. 6). This worry should give experimental biologists wishing to employ MM some reason to pause. Craver (2007, p. 158) worried, for example, that MM might allow the heart to qualify as a component in the mechanism underlying performance in word-stem completion tasks, even though it seems to be more appropriately described as a causal background condition for that performance. Lesioning or stimulating the heart in a bottom-up experiment would almost certainly disrupt a subject’s ability to perform word-stem completion. One could imagine that by engaging a subject in an especially arduous version of a word-stem completion task, the heart might start beating faster. According to MM, then, the heart would count as a component.

Craver suggests that in cases like this we might appeal to the additional fact that interventions on background conditions as opposed to components will tend to be relatively nonspecific in their effects. Lesioning or stimulating the heart, to the extent that it has any effects at all, will likely have diffuse effects that disrupt or compromise performance in the target task of word-stem completion and many other tasks besides. Relatedly, interventions on background conditions will tend not to have subtle, differential effects on word-stem completion (e.g., error rates are unlikely to vary parametrically with lesion size or stimulation intensity). Instead, these interventions will tend to have unsubtle, all-or-none effects (e.g., complete extinguishing of task performance).

This idea of specificity has been subject to considerable discussion under the rubric of *causal specificity* (Griffiths et al., 2015; Waters, 2007; Woodward, 2010). Woodward (2010) used a simple example to help illustrate the basic idea. He asked us to consider the difference between the tuning dial and the on/off switch on a radio. Both are capable of exerting a causal influence on what the listener hears: The dial must be appropriately tuned to a specific frequency (e.g., 105.7 MHz), and the on/off switch must be in the “on” state. But important to note, the dial and switch seem to differ with respect to the degree to which they are “specific” to their effects. The tuning dial has relatively fine-grained causal influence over which station is heard by the listener (assuming the switch is on) in the sense that there are many possible values the dial can take (the cause variable), which result in correspondingly many differences in the station that is heard (the effect variable). The dial is thus a relatively *specific* cause of which station is heard.

By contrast, although the on/off switch is also causally relevant to whether any station is heard at all, it has no causally specific influence on which particular station is heard. The switch is only a relatively *nonspecific* cause of whether a station is heard.

This qualitative distinction between specific and nonspecific causes has recently been made more precise by applying tools from information theory. Griffiths et al. (2015) suggested that causal specificity can be measured in terms of the reduction in uncertainty about the value of the effect variable E that results from intervening to set the value of the cause variable C (i.e., the *mutual information* between E and C).

How, then, might this distinction between dial-like or specific causes, on one hand, and switchlike or nonspecific causes, on the other, be leveraged for thinking about extended components in cognition? Consider the interesting case of extended spider cognition discussed at some length by Cheng. The experimental results seem to show that MM is satisfied. Specifically, experimental interventions to increase thread tension in particular areas of the web change spider foraging behavior. In the reverse direction, interventions to induce changes in the internal state of the foraging system (changes in spider satiation levels) can lead to changes in web thread tension. But we may also want to ask additional questions about the specificity of these effects. For example, do fine-grained changes in web thread tension change spider foraging behavior in specific or nonspecific ways? Put another way, is web thread tension more dial-like or more switchlike in its influence on spider behavior? There is some evidence that relatively fine-grained changes in web tension can have similarly fine-grained or differential effects on spider foraging behavior and attentional state (Japyassú & Laland, 2017; Nakata and Zschokke, 2010; Watanabe, 2000). In the other direction, do fine-grained changes in spider satiation levels change web thread tension in specific or nonspecific ways? At least one experimental study seems to suggest that they do (Japyassú & Laland, 2017; Watanabe, 2000).

An interesting next step for investigations of extended cognition like these might be to compare the specificity “profiles” of internal versus external components using well-defined information-theoretic measures like mutual information (Griffiths et al., 2015). Because interventions on typical (i.e., internal) components will tend to produce relatively specific downstream effects as compared to those on causal background conditions, might the same be true for external components? If interventions on external components are equally specific in

their effects, this could provide an additional line of support for situated cognition. If their specificity profiles turned out to be systematically different in some way, what would this mean?

We pose these questions not because we have answers but because we strongly suspect that by addressing these and other related questions, research on embodied and extended cognition in nonhuman animals might be propelled into some fruitful and unexpected directions.

References

- Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47, 139–159. doi:10.1016/0004-3702(91)90053-M
- Brooks, R. A. (1999). *Cambrian intelligence*. Cambridge, MA: MIT Press.
- Clark, A. (1997). *Being there*. Cambridge, MA: MIT Press.
- Clark, A. (2005). Word, niche and super-niche: How language makes minds matter more. *Theoria*, 20, 255–268.
- Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. New York, NY: Oxford University Press.
- Flash, T., & Sejnowski, T. J. (2001). Computational approaches to motor control. *Current Opinion in Neurobiology*, 11, 655–662. doi:10.1016/S0959-4388(01)00265-3
- Griffiths, P. E., Pocheville, A., Calcott, B., Stotz, K., Kim, H., & Knight, R. (2015). Measuring causal specificity. *Philosophy of science*, 82, 529–555. doi:10.1086/682914
- Hauser, H., Ijspeert, A. J., Fuchslin, R. M., Pfeifer, R., & Maass, W. (2011). Towards a theoretical foundation for morphological computation with compliant bodies. *Biological Cybernetics*, 105, 355–370. doi:10.1007/s00422-012-0471-0
- Hauser, H., Ijspeert, A. J., Fuchslin, R. M., Pfeifer, R., & Maass, W. (2012). The role of feedback in morphological computation with compliant bodies. *Biological Cybernetics*, 106, 595–613. doi:10.1007/s00422-012-0516-4

- Hoffmann, M., & Müller, V. C. (2017). Simple or complex bodies? Trade-offs in exploiting body morphology for control. In G. Dodig-Crnkovic & R. Giovagnoli (Eds.), *Representation and reality in humans, other living organisms and intelligent machines* (pp. 335–345). New York, NY: Springer. doi:10.1007/978-3-319-43784-2_17
- Hutchins, E. (2010). Cognitive ecology. *Topics in Cognitive Science*, 2, 705–715. doi:10.1111/j.1756-8765.2010.01089.x
- Japyassú, H. F. (2017). Plasticity and cognition in spiders. In C. Viera & M. O. Gonzaga (Eds.), *Behaviour and ecology of spiders* (pp. 383–396). New York, NY: Springer. doi:10.1007/978-3-319-65717-2_14
- Japyassú, H. F., & Laland, K. N. (2017). Extended spider cognition. *Animal Cognition*, 20, 375–395. doi:10.1007/s10071-017-1069-7
- Kaplan, D. M. (2012). How to demarcate the boundaries of cognition. *Biology and Philosophy*, 27, 545–570. doi:10.1007/s10539-012-9308-4
- Kirsh, D. (1991). Today the earwig, tomorrow man? *Artificial Intelligence*, 47, 161–184. doi:10.1016/0004-3702(91)90054-N
- Menary, R. (2007). *Cognitive integration: Mind and cognition unbounded*. New York, NY: Springer. doi:10.1057/9780230592889
- Nakajima, K., Hauser, H., Kang, R., Guglielmino, E., Caldwell, D. G., & Pfeifer, R. (2013). A soft body as a reservoir: Case studies in a dynamic model of octopus-inspired soft robotic arm. *Frontiers in Computational Neuroscience*, 7. doi:10.3389/fncom.2013.00091
- Nakata, K., & Zschokke, S. (2010). Upside-down spiders build upside-down orb webs: Web asymmetry, spider orientation and running speed in *Cyclosa*. *Proceedings of the Royal Society of London B: Biological Sciences*, 277(1696). doi:10.1098/rspb.2010.0729
- Pfeifer, R., & Bongard, J. (2006). *How the body shapes the way we think: A new view of intelligence*. Cambridge, MA: MIT Press.
- Pfeifer, R., Iida, F., & Lungarella, M. (2014). Cognition from the bottom up: On biological inspiration, body morphology, and soft materials. *Trends in Cognitive Sciences*, 18, 404–413. doi:10.1016/j.tics.2014.04.004
- Pfeifer, R., Lungarella, M., & Iida, F. (2007). Self-organization, embodiment, and biologically inspired robots. *Science*, 318(5853), 1088–1093. doi:10.1126/science.1145803
- Sumbre, G., Fiorito, G., Flash, T., & Hochner, B. (2006). Octopuses use a human-like strategy to control precise point-to-point arm movements. *Current Biology*, 16, 767–772. doi:10.1016/j.cub.2006.02.069
- Sutton, J. (2015). Remembering as public practice: Wittgenstein, memory, and distributed cognitive ecologies. In *Mind, Language & Action: Proceedings of the 36th International Wittgenstein Symposium* (pp. 409–443). Berlin, Germany: De Gruyter. doi:10.1515/9783110378795.409
- Watanabe, T. (2000). Web tuning of an orb-web spider, *Octonoba sybotides*, regulates prey-catching behaviour. *Proceedings of the Royal Society of London B: Biological Sciences*, 267(1443), 565–569. doi:10.1098/rspb.2000.1038
- Waters, C. K. (2007). Causes that make a difference. *The Journal of Philosophy*, 104, 551–579. doi:10.5840/jphil2007104111
- Webb, B. (1994). Robotic experiments in cricket phonotaxis. In D. Cliff, P. Husbands, J.-A. Meyer, & S. W. Wilson (Eds.), *From animals to animats 3: Proceedings of the Third International Conference on Simulation of Adaptive Behavior* (pp. 45–54). Cambridge, MA: MIT Press.
- Wilson, R. A. (1994). Wide computationalism. *Mind*, 103, 351–372. doi:10.1093/mind/103.411.351
- Woodward, J. (2010). Causation in biology: Stability, specificity, and the choice of levels of explanation. *Biology & Philosophy*, 25, 287–318. doi:10.1007/s10539-010-9200-z