

VEHICLES, CONTENTS, CONCEPTUAL STRUCTURE, AND EXTERNALISM

S. L. Hurley

Analysis 1998, vol. 58, no. 1, pp. 1-6

We all know about the vehicle/content distinction (see Dennett 1991a, Millikan 1991, 1993). We shouldn't confuse properties represented in content with properties of vehicles of content. In particular, we shouldn't confuse the personal and subpersonal levels. The contents of the mental states of subject/agents are at the personal level. Vehicles of content are causally explanatory subpersonal events or processes or states. We shouldn't suppose that the properties of vehicles must be projected into what they represent for subject/agents, or vice versa. This would be to confuse the personal and subpersonal levels.

Consider the generality and systematic conceptual structure of the cognitive abilities of persons. Content at the personal level has decompositional, recombinant conceptual structure. What explains this structure at the level of subpersonal processing? What properties of vehicles of content make for conceptually structured content? Do subpersonal vehicles of conceptual content themselves have a language-like or syntactic structure, so that structure in personal-level content is explained by isomorphic subpersonal structure? Their answers to this question distinguish classical, 'language of thought' approaches to cognitive architecture from connectionist approaches. Classicists answer *yes*, on either a priori or empirical grounds, while connectionists allow that it may be possible to explain conceptually structured cognitive abilities in terms of neural networks without syntactic

structure. The assumption that the processes that support true thought must have a classical architecture (even if they are implemented by a connectionist network) imposes a requirement of *causal systematicity* on thought.¹ On its face, this looks like a vehicle/content confusion.

But might our conception of ourselves as genuine thinkers nevertheless impose this particular projection from contents to vehicles? If the conceptual structure of content were not explained by isomorphic subpersonal structure, would displays of 'cognitive ability' be mere patterns of behaviour, mere mimicry of true thought? This is a kind of zombie or marionette worry, about thought rather than consciousness. So some argue that if connectionism turns out to give the right account of supposedly cognitive abilities, then thought is eliminated.

Much here turns on two questions. First, should the requirement of causal systematicity be applied, if at all, to what explains *particular* thoughts and particular displays of cognitive ability, or to what explains why any thought content or cognitive ability of a given *type* exists at all? Vehicles of content explain tokens, particular thoughts on particular occasions. That is, they are *token-explanatory* processes, not merely processes that explain why a certain type of content exists at all. It is the application of a requirement of causal systematicity to token-explanatory processes that looks like a vehicle/content confusion. But perhaps causal systematicity should be applied instead to *type-explanatory* processes. This might capture what is intuitively right about requiring causal systematicity while avoiding any vehicle/content confusion.

Second, must what explains conceptually structured cognitive abilities for purposes of causal systematicity be *internal*? Or can it be world-involving? Can causal systematicity be satisfied relationally, by processes or states that cross the boundary between organism and environment? If it can be, does it matter for various purposes whether it is also satisfied internally?²

This second question takes two forms, as applied to type-explanatory and to token-explanatory processes. The answer to the first question is obvious. Type-explanatory processes can be world-involving: consider upbringing, evolution, and so on. The answer to the second question is less obvious. Can token-explanatory processes, or vehicles of content, also be world-involving processes or relational states of persons? A full explanation of a token thought would include type-explanatory processes, which explain why contents of that type exist at all. But tacking external type-explanatory processes such as upbringing or evolution onto a token-explanatory process is not enough to make for a partly external vehicle. The vehicle is bounded by what's left when you subtract the type-explanatory processes from the full explanation of the token. The question is whether *that* could go external or relational. Arguments both for and against externalism about content typically presuppose internalism about vehicles (so that internal *physical* states can be duplicated despite environmental inversions, even if their content supervenes on external relations). But is the presupposition of vehicle internalism justified?

There is nothing spooky about relational properties of persons: they correspond to intrinsic properties of something bigger where causality reigns as usual. Context-dependence is relative to a boundary. We are used to the idea of relations internal to a

nervous system as vehicles: relations between events in a distributed process within the brain need not be monitored by a single cell or module. Such internal relations may themselves be the vehicles of content, rather than the intrinsic properties of some internal monitor of those relations. But if internal relations can qualify as vehicles, why not external relations? Given a continuous complex dynamic system of reciprocal causal relations between organism and environment, what in principle stops the spread? The idea that vehicles might go external takes the notion of distributed processing to its logical extreme.

The question whether vehicles can go external is raised but not answered here. But a bad answer is that they cannot simply because there can be thoughts about a thing on particular occasions when the thinker is not in contact with it, when it is absent from her environment. On such occasions, the vehicle does not involve the worldly thing, since the vehicle explains the token thought and the worldly thing it is about is absent. But we should not assume the vehicle of such a thought must necessarily be the same as the vehicle of a thought about the same thing when the thinker is in contact with it. That would be to project sameness from contents to vehicles. For example, the vehicles of geometrical reasoning about objects that are perceptually presented on a particular occasion need not be the same (though may be) as the vehicles of similar reasoning in imagination. If vehicles can go external, the former might include pen and paper, while the latter do not.

Friends of causal systematicity tend to assume it should be applied to internal, token-explanatory vehicles of content. This assumption appears to involve a vehicle/content confusion in the ways just noted. Can the intuition that genuine thought

requires causal systematicity be captured instead by applying it to external, type-explanatory processes?

Compare two subject/agents, Mycroft and Marlowe, and consider whether they are genuine thinkers. Both seem to display conceptually structured cognitive abilities. They speak and understand, read and write, reason and argue logically. But is this mere mimicry of true thought? Internally, they look rather different. Inside Mycroft we find syntactically structured token-explanatory vehicles of content. Inside Marlowe we find only an elaborate neural network that implements no syntactic structure.

But now zoom back to enlarge the picture, and consider why they have the type of cognitive abilities they have at all. Mycroft is a human being, not a machine. But his brain was programmed by a Martian philosopher who abandoned him when she no longer had a use for him. Most of his 'thinking' is done from his armchair. On the other hand, Marlowe's network was trained in part by a linguistic upbringing and education and is continuously tuned by his robust sensory-motor interactions with his environment, including natural language (such as conversations). Against an evolutionary background, the interaction of Marlowe's neural network with environments including natural language explains the existence of his cognitive and logical abilities. And of course natural language has syntactic structure, generality, systematicity, and so on.

Does it matter that if we truncate the explanation of Marlowe's cognitive abilities and look only internally and only at token-explanatory processes we find no reduplication of syntactic structure inside Marlowe? Intuitively, no. Marlowe's title to genuine thought is

more secure than Mycroft's, anchored precisely by his ongoing interactions with his environment. If either, it is Mycroft who is the mere mimicker of thought, the appropriate object of zombie or marionette worries.³

Arguments from the success of connectionism to the elimination of thought tend to assume that, to do the explanatory work needed for true thought, syntactic structure must be located internally: in the head as well as in the system that includes both head and environment with which we interact. But this is intuitively wrong, as Mycroft and Marlowe show. If true thought requires causal systematicity at all (which is not clear), the requirement can be met by type-explanatory processes and relationally, in the way just illustrated. Causal systematicity can be satisfied by the whole system, including embedding linguistic environment, when it is not satisfied by internal vehicles.

A view of perception and action as separate input and output systems often goes with a view of thought and cognition as 'central' and in turn separate from the 'peripheral' input and output systems. On this view, the virtual processing of cognition is central, even if its implementation is distributed. Input to it is provided by perception, and it issues output that generates action. Against this background, a standard view is that connectionist approaches are at their strongest for the peripheral processes (such as sensorimotor control, pattern recognition), but that the central cognitive interface must have classical structure. The mind is seen as a kind of sandwich in which perception and action hold a classical center in place.

This picture can be attacked at the center, by arguing against the view that true thought requires causal systematicity for internal vehicles of thought. One such argument is given above. But attacks at the centre too often accept the conception of perception and action as peripheral buffer zones, which holds the central interface conception of cognition in place. Another way to undermine the classical sandwich is to attack from the outside, as it were: to challenge the buffer zone view and the separation of perception and action from central cognitive processes. The two angles of attack are complementary and mutually supporting.

We've seen how zombie worries about 'cognitive abilities' without explanatory, classically structured internal vehicles can trade on a spurious dichotomy: either the right explanatory internal structure, or mere mimicry, mere patterns of behaviour. The conception of perception and action in terms of input and output, respectively, encourages this dichotomy by type-casting behaviour, including linguistic behaviour, in the role of effect and ignoring its other talents. As the mere effect of thought, behaviour is merely evidence for thought. So worries about behaviourism and verificationism get started.

But linguistic behaviour, like other behaviour, is as much cause as effect within a complex dynamic feedback system. And complex dynamic systems can have very surprising emergent properties. Indeed, emergence in complex dynamic systems could be argued to underlie and explain the vehicle-content distinction. An essential lesson we learn from studying nonlinear systems is that discontinuities and structure at a higher level can emerge unpredictably (though deterministically), with no isomorphism at the underlying

level (see Elman et al 1996, Kelso 1995, Port and van Gelder 1995, Thelen and Smith 1994, etc.).

To map the distinction between perception and action and the distinction between input and output onto one another itself looks like another confusion of personal and subpersonal levels. A better conception of perception and action is as co-dependent on a complex dynamic system of causal relations, a system that includes subpersonal relations between inputs to and outputs from the organism but that may also extend into the environment.⁴ This improvement can underscore the obvious but neglected point that to participate in natural language is to perceive and to act, and depends on sensory-motor interactions with linguistic environments. It can also help to understand how the conceptual structure of cognitive abilities might emerge from such interactions, even in the absence of internal causal systematicity.⁵

Consider a connectionist network trained and tuned in part by systematic sensory-motor interactions with a richly structured linguistic environment, as feedback loops pass through this environment picking up information. Functions from the network's linguistic output back to linguistic input reflect linguistic structure in the environment and constrain the network's organization. For example, organization in the network could be induced in part by the way self-produced words result in other-produced words. Such organization might permit new combinations of words, answers to new questions: networks can generalize beyond the cases they are trained on. Behaviour does causal work in such a system; it's not merely evidence of something else that's doing the work.

It is a further question why there is natural language in the environment at all. Of course, not just any old internal architecture makes natural language possible. But it is an empirical question whether whatever explains the existence of natural language involves internal syntactic structure. Our conception of ourselves as thinkers does not require this to be the case.⁶

University of Warwick

Coventry CV4 7AL, UK

REFERENCES

- Clark, A. 1990. *Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing*. Cambridge: The MIT Press.
- Davies, M. 1991. Concepts, connectionism and the language of thought. In *Philosophy and Connectionist Theory*, ed. W. Ramsey, S. P. Stich, and D. E. Rumelhardt, 229-57. Hillsdale, New Jersey: Erlbaum.
- Dennett, D. C. 1991a. *Consciousness Explained*. Boston: Little Brown and Company.
- Dennett, D. C. 1991b. Mother Nature versus the walking encyclopedia: a Western drama. In *Philosophy and Connectionist Theory*, ed. W. Ramsey, S. P. Stich, and D. E. Rumelhart, 21-30. Hillsdale, New Jersey: Erlbaum.
- Elman, J, et al. 1996. *Rethinking Innateness*. Cambridge: MIT Press.
- Hurley, S. L., forthcoming 1998. *Consciousness in Action*. Cambridge: Harvard University Press.
- Hutchins, E. 1995. *Cognition in the Wild*. Cambridge: MIT Press.
- Kelso, J. A. S. 1995. *Dynamic Patterns*. Cambridge: MIT Press.
- Millikan, R. G. 1991. Perceptual content and Fregean myth. *Mind* 100:439-59.

Millikan, R. G. 1993. Content and vehicle. In *Spatial Representation*, ed. N. Eilan, R. McCarthy, and B. Brewer, 256-68. Oxford: Blackwell.

Port, R. F. and van Gelder, T., eds. 1995. *Mind as Motion*. Cambridge: MIT Press.

Thelen, E. and Smith, L. B. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge: MIT Press.

NOTES

1. Cf. Davies 1991, p. 246: 'Mirroring the commonality in the inferences that are found primitively compelling--namely, their form--there should be a commonality in the causal processes that explain their being so.' --But in the truncated internal portion of these processes?
2. Cf. Clark 1990:64; Dennett 1991b:26, on 'sentences lying around'; Hutchins 1995.
3. I am indebted to Nicholas Rawlins for casting Mycroft and Marlowe in this illustration. For those who were never Sherlock Holmes fans, Mycroft is the famous detective's still more cerebral but inactive older brother.
4. This is argued for in Hurley, forthcoming, which also discusses in more detail many of the issues raised briefly in this article.
5. Appeals to structured information in the environment have a natural affinity with connectionism: consider ecological approaches to perception for another example.
6. For discussion of these issues I am grateful to Ron Chrisley, Andy Clark, Michael Wheeler, Sonia Sedivy, Nicholas Rawlins, and to members of the audiences at my talks in UCSD and UCSB in October 1996 when I presented this material. I am also grateful to the British Academy and to the McDonnell-Pew Centre for Cognitive Neuroscience for support of this work.