

Letter

Animals and the iterative natural kind strategy

Tim Bayne ^{1,2,*}, Anil Seth^{2,3},
 Marcello Massimini^{2,4},
 Joshua Shepherd^{2,5},
 Axel Cleeremans^{2,6},
 Stephen M. Fleming^{2,7,8},
 Rafael Malach^{2,9},
 Jason B. Mattingley^{2,10},
 David K. Menon^{2,11},
 Adrian M. Owen^{2,12},
 Megan A.K. Peters^{2,13},
 Adeel Razi^{2,8,14}, and
 Liad Mudrik^{2,15}

What entities are conscious and how could we ever know? We recently provided a framework for developing validated tests for consciousness ('C-tests'): the Iterative Natural Kind Strategy (INKS) [1]. Brown and colleagues [2] raise some important questions for the INKS approach, although we are not sure how different their answers to those questions are from our own.

'To test the boundaries of consciousness, study animals', the title of their letter reads. We could not agree more. If there is a contrast between their approach and ours, it is surely not about whether non-human animals should be studied, but about the role that their study ought to play in this project.

Much of their concern focuses on Figure 2 in [1], reprinted here (Figure 1). This figure, they claim, 'gives the impression that non-human animals (especially invertebrates) are amongst the most distant populations, with most tests deemed easier to apply and interpret in humans undergoing disordered states of consciousness. This implies that we should start by investigating [artificial intelligence] AI and coma patients before eventually turning our attention to

other active, mobile animals with brains. That would be a massive misstep [2].

It would indeed be a mistake to leave non-human animals to last, but it is not a mistake we make. We refrained from offering an account of how populations ought to be stratified into levels (i.e., distance from 'consensus cases' – roughly, neurotypical adult humans), and we certainly did not claim that non-human animals are furthest from the consensus cases. On the contrary, we explicitly grouped non-human animals that are 'close relatives of ours' together with non-responsive humans and human infants (and fetuses) as 'neighboring populations' we also described organoids, xenobots, and AI systems as 'clearly less close to the consensus cases' than all other populations.

Brown and colleagues seem to have taken this figure as an attempt to capture (and justify) the stratification of populations into levels. That was not its point. It was not primarily concerned with what a population might contribute to the validation of a C-test, but with what C-tests can tell us about consciousness in that population. Can a C-test even be applied to the relevant population? If so, what verdict might it render? Brown *et al.* worry that many of the claims embedded in this figure are contestable and that we ought to have justified them. Indeed, the claims are contestable; this is unavoidable given the uncertainties involved. Our aim was not to render a definitive verdict on consciousness in any particular population, but to take some initial steps towards a large and complex project to which we hope others will contribute. We are pleased that Brown *et al.* are doing just that.

But perhaps the most fundamental issue here is whether the INKS approach to validation should be 'hierarchical'. An approach is hierarchical if it treats certain populations ('levels') as privileged relative to others with respect to validating a C-

test. In [1] we embraced a hierarchical approach, suggesting that validation ought to begin with ordinary adult humans (the 'consensus population') before being extended to other populations in an iterative fashion. What attitude do Brown *et al.* take to this issue?

On one reading of their letter, they reject any distinction between populations, holding that even judgements about consciousness in healthy adult humans have no evidential priority. That view would be at odds with current practices; indeed, each of the C-tests mentioned by Brown and colleagues was first validated in humans. For example, it is only thanks to dream reports provided by humans that we have learned to associate rapid-eye movement (REM)-like sleep and measures related to eye movements with the presence of conscious experience. A more moderate reading accepts that judgments about consciousness in healthy adult humans ('level 0') have evidential priority, but refuses to make distinctions between populations beyond that (i.e., everything else is 'level 1'). The problem with this position, we would argue, is that the rationale for distinguishing between two levels seems to generalize. As Brown *et al.* note, it seems implausible to treat data concerning AI on a par with data concerning non-human animals,

Indeed, despite describing the hierarchical approach as 'controversial', they themselves seem to presuppose it, for they say that the first phase of validation ought to include 'a wide range of non-human animals', implying that some animals need not be excluded. But which animals make the cut? There are difficult questions here for all of us, unless, of course, one assumes that all animals (even Placozoa?) are conscious [3], an assumption that few would want to make.

We have focused on the (apparent) contrasts between our views and those of

		Command following	Narrative	Sniffing	PCI	Global effect	ACT	UAL
ALTERED STATES	Sedation	+	+	+	+	+	?	+
	Epileptic seizure	+	+	+	+	+	?	+
	Sleep/Dreaming	+ ?	+ ?	+ ?	+	+ ?	?	+ ?
UNCLEAR CAPACITY FOR CONSCIOUSNESS	Disorders of consciousness	+	+	+	+	+	?	+
	Babies	-	?	+	+	+	-	+
	Fetuses	-	-	?	+	+	-	?
NON-HUMAN ANIMALS	Non-human mammals	-	-	+	+ ?	+	-	+
	Non-mammal vertebrates	-	-	-	-	+	-	+
	Invertebrates	-	-	-	-	?	-	+
ARTIFICIAL SYSTEMS	Neural organoids	-	-	-	?	+ ?	-	+ ?
	xenobots	-	-	-	?	+ ?	-	+ ?
	AI	+ ?	+ ?	-	?	+ ?	+	+ ?

Trends in Cognitive Sciences

Figure 1. The scope of tests for consciousness (C-tests). The applicability of different C-tests (rows) to different populations (columns), divided into levels based on the provisional order of validation suggested here. Plus (+) signs indicate that a test can likely be administered to a specific population in a meaningful way (although its specificity/ sensitivity might be low), possibly with some modifications. Dashes (-) denote the inapplicability or irrelevance of the test for a specific population. Question marks (?) denote that the test might be applicable, but more development is needed to test whether this is the case. Finally, a combination of a plus sign and a question mark (+?) signifies that although the test can be applied it is unclear what its results would mean. Reproduced from [1]. Abbreviations: ACT, AI consciousness test; AI, artificial intelligence; PCI, perturbational complexity index; UAL, unlimited associative learning.

Brown *et al.*, but it is important to recognize that on the fundamental questions that confront us in this space (Is it possible to make significant progress in developing validated C-tests? Does validation require treating consciousness as a natural kind?), we are very much on the same page.

Acknowledgments

The authors gratefully acknowledge support from the 'Brain, Mind and Consciousness' program of the Canadian Institute for Advanced Research (CIFAR).

Declaration of interests

M.M. is cofounder and shareholder of Intrinsic Powers.

- ³Centre for Consciousness Science and School of Engineering and Informatics, University of Sussex, Sussex, UK
- ⁴University of Milan, Milan, Italy
- ⁵ICREA/Universitat Autònoma de Barcelona, Barcelona, Spain
- ⁶Center for Research in Cognition & Neuroscience, ULB Institute of Neuroscience, Université libre de Bruxelles, Brussels, Belgium
- ⁷Department of Experimental Psychology, University College London, London, UK
- ⁸Wellcome Centre for Human Neuroimaging, University College London, London, UK
- ⁹The Weizmann Institute, Rehovot, Israel
- ¹⁰Queensland Brain Institute & School of Psychology, The University of Queensland, St Lucia, Queensland, Australia
- ¹¹Cambridge University, Cambridge, UK
- ¹²University of Western Ontario, Ontario, Canada
- ¹³University of California Irvine, Irvine, CA, USA
- ¹⁴Turner Institute for Brain and Mental Health, Monash University, Melbourne, Victoria, Australia
- ¹⁵School of Psychological Sciences and Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel

*Correspondence: timothy.bayne@monash.edu (T. Bayne). <https://doi.org/10.1016/j.tics.2024.07.009>

© 2024 Published by Elsevier Ltd.

References

1. Bayne, T. *et al.* (2024) Tests for consciousness in humans and beyond. *Trends Cogn. Sci.* 28, 454–466
2. Brown, S. *et al.* (2024) To test the boundaries of consciousness, study animals. *Trends Cogn. Sci.*, (in press)
3. Andrews, K. (2024) "All animals are conscious": shifting the null hypothesis in consciousness science. *Mind Lang.* 39, 415–433