Daniel Dennett

It's Not a Bug, It's a Feature

Today, the planet has plenty of conscious beings on it; three billion years ago, it had none.[1] What happened in the interim was a lot of evolution, with features emerging gradually, in one order or another. Figuring out what order and why is very likely a good way to reduce perplexity, because one thing we have learned from the voyage of the Beagle and its magnificent wake is that puzzling features of contemporary phenomena often are fossil traces of earlier adaptations. As the great biologist D'Arcy Thompson once said, 'Everything is the way it is because it got that way' And even when we can't remotely confirm our Just So Stories about how things got the way they are, the exercise can be salutary, since it forces us to ask (and try to answer) questions that might otherwise never occur to us. We do have to get the left and right sides of our equation to match in dimensionality — I am grateful to Humphrey (2000)for this useful proposal about how to think about the issues — and adding wrinkles on the right needs to be motivated by, and in the end justified by, more than the sheer need for a few more dimensions. As Just So Stories go, Humphrey's account of the emergence of sensation is a valuable one, traversing ground that must be traversed one way or another, and providing along the way some reasonable grounds for supposing things happened roughly the way he supposes.

Humphrey has convinced me that *something like* his distinction between visual sensation and visual perception needs to be drawn, but rather than focus on relatively minor problems I have with specifics of his account, I want to articulate and then rebut a blanket 'objection' that I anticipate will be widespread in other commentaries on this essay:

> A robot could meet all of Humphrey's dimensional conditions. Yes, of course. Humphrey frames the design of his conscious organism in terms of evolutionary redesign, and stresses the ecological interplay that helps set the costs and benefits for this exercise in R-and-D, but nothing he proposes in the way of an evolutionary innovation is *in principle* beyond the reach of roboticists. For instance, he says at a midway point in his Just So Story: '. . . the animal is actively responding to stimulation with public bodily activity, and its experience or proto-experience of sensation (if we can now call it that) arises from its monitoring of its own command signals for these sensory responses' I

[1] You will agree unless you are one of those who wants to grant consciousness to bacteria and other single-celled life forms. Granting a smidgen — or perhaps a 'quantum' — of micro-consciousness to bacteria is a logically available option, with nothing to recommend it and many problems, as I explain elsewhere (Dennett, forthcoming).

Journal of Consciousness Studies, **7**, No. 4, 2000, pp. 25–7

> presume that a robot can 'actively' respond and is capable of at least '*proto*-experience of sensation'; if these presumptions are not so, Humphrey is smuggling in something crucial with these terms. So Humphrey is, in spite of his assurances, only dealing with the easy problems of consciousness, since even if he is right about everything he says, he has provided an account only of those features of consciousness that are robot-friendly, functionalistic, a matter of 'complex behavioral dispositions' — and, as he says of Dennett's earlier attempt, such an account, 'while defensible in his own terms, has proved too far removed from most people's intuitions to be persuasive.

I think the correct response to this objection is as follows: Yes, indeed, in principle a robot *could* instantiate Humphrey's theory. But not just *any* robot. It would have to be a robot quite unlike the typical robots of both reality and imagination, and whether or not it could *actually* be created is an empirical question. (A conscious robot, like a splittable atom, may be held to be 'impossible by definition' — but definitions can go extinct when they've outlived their usefulness.) Humphrey makes an important point when he claims that our sensory states are descendants of more primitive earlier systems of response-to-stimulation, and as such come already linked quite tightly to action-propensities that can be suppressed or deflected only by mounting layers of competing forces and coalitions; additional structures that modify the settings and import of the ancestral types, while preserving their evaluative valence. So we'd have to permit the roboticists to give their robot a *virtual* past, with pain-wiggles and salt-wiggles and the like, leaving their fossil traces on the (hand-coded, not naturally selected) designs of the 'descendant' systems. It would have to be a robot with a particular sort of organization, the sort of organization that *might* be artificially created but that would arise naturally by something like the process described in Humphrey's Just So Story. It would have to be an embodied robot, like Cog (Dennett, 1998, chapter 9).[2] Its nano-machinery would not necessarily have to be protein molecules (like ours), but it would display both the functions *and dysfunctions* that we display, thanks to our evolutionary heritage. For instance, it would find some topics harder to concentrate on than others simply because the sensory baggage that those topics carried was, for 'prehistorical' reasons, harder to overcome. A trivial example: it wouldn't just show human performance deficits on the Stroop test (reading colour names printed in ink of non-matching colours); it would prefer red ink for some topics and green ink for others, for reasons it found impossible to articulate. Multiply this case by a thousand. In every circumstance in which people manifest — and *sometimes* reflect on — such differential loading (was the element Humphrey calls sensation present or not, and if present, what was its evaluative valence, if any?) the robot would do likewise because it, too, was endowed with an organization having the strengths and concomitant weaknesses provided by such an evolutionary history.

Now the question to consider is whether a robot that matched human function and dysfunction at such a grain level would be conscious. If you are sure that the answer is *no*, you should reflect on what your reason could possibly be, given the deliberate sketchiness of the foregoing description. If your reason is only that you insist on maintaining a vision of consciousness that is *automatically* proof against any kind of robot, you are just retroactively adding dimensions — one might suspect: *making up*

2] And yes, it is only 'practical' considerations that demand this: 'in principle' it could live its whole life as a brain in a vat, though the vat would have to be Vast [Dennett, 1995, p. 109] in its complexity in order to provide the full force of *virtual* embodiment.

dimensions — to put on the left hand side of the equation. In sum, the fact that Humphrey's account leaves open the prospect of a conscious robot is in its favour, not a problem. As they say in the software world, 'It's not a bug, it's a feature'

References

Dennett, D.C. (1995), *Darwin's Dangerous Idea* (New York: Simon & Schuster).
Dennett, D.C. (1998), *Brainchildren* (Cambridge, MA: MIT Press).
Dennett, D.C. (forthcoming), 'The Zombic Hunch: The extinction of an intuition?' *Philosophy*.
Humphrey, N. (2000), 'How to solve the mind–body problem' *Journal of Consciousness Studies*, 7 (4), pp. 5–20 (this issue).

Published in the UK by Imprint Academic
PO Box 1, Thorverton EX5 5YX, UK

Published in the USA by Imprint Academic
Philosophy Documentation Center, Bowling Green State Universi
Bowling Green, OH 43403-0189, USA

ISBN 0 907845 08 8 (paperback)

ISSN 1355 8250 (*Journal of Consciousness Studies*, **7**, No.4, 2000)

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library
Library of Congress Card Number : 00–101880

Cover illustration: Claire Harper

Printed in Exeter UK by Short Run Press Ltd