Richer, J. (2021). The Evolution of Consciousness - the Shortcomings of: *The Ancient Origins of Consciousness* by Todd Feinberg and Jon Mallatt, and an Alternative Approach. Human Ethology, 36, 27-35. https://doi.org/10.22330/he/36/027-035

Human Ethology 36(2021): 027-035
Commentary

THE EVOLUTION OF CONSCIOUSNESS – THE SHORTCOMINGS OF: *THE ANCIENT ORIGINS OF CONSCIOUSNESS* BY TODD FEINBERG AND JON MALLATT, AND AN ALTERNATIVE APPROACH

John Richer

Department of Physiology, Anatomy and Genetics, University of Oxford, UK.
Paediatric Psychology, Oxford University Hospitals NHS Trust, UK.

johnricher@oxhs.co.uk

ABSTRACT

The scientific study of consciousness is a a seductive topic given that we are a species which ascribes consciousness to each other. We communicate on the assumption others are conscious agents like ourselves. So to investigate the evolution of this phenomenon as in the book *The Ancient Origins of Consciousness*[1] would seem promising. Unfortunately, despite great detail and eruditon in the discussion of evolution, the endeavour hits the same obstacles and make the same mistakes as is seen in swathes of psychology. The central error is that consciousness is treated as the same sort of publicly observable phenomenon as behaviour or brain physiology and so can be the subject of scientific study. But it isn't and it can't, and this paper offers an explaination why.

[1] The Ancient Origins of Consciousness
By Todd Feinberg and Jon Mallatt, 2016
MIT Press. Cambridge
ISBN 978-0-262-53460-4

submitted: 26. Jan. 2021; revised: 29. Jan. 2021; accepted: 30.Jan. 2021

"THE ANCIENT ORIGINS OF CONSCIOUSNESS"

Any ethological study of human behaviour needs to take into account that we are a species which communicates intersubjectively, embracing the mental states and the first person perspective of others, treating them as conscious beings like ourselves. Therefore a book which discusses consciousness in the context of evolution and biology looks as if it might address some key issues of interest to human ethologists. With this hope in mind I read *The Ancient Origins of Consciousness* by Todd Feinberg and Jon Mallatt, a Psychiatrist and Zoologist respectively

The authors certainly flag up early that there are philosophical issues about whether consciousness can be studied scientifically. They quote the philosopher John Searle asserting that it is impossible to reduce the first person (i.e. subjective,) experience to patterns of neural firing, or vice versa. Towards the end they quote Gordon Globus (1973) on the difference between subjective and objective descriptions; from the subject's viewpoint, neural events are not themselves experienced the subject has only the experience itself, whereas from the objective viewpoint neural events can (theoretically) be observed but the experience of the subject cannot. The former impossibility (they call it an irreducibility), the authors call "auto ontological irreducibility" and the later "allo ontological irreducibility".

Having early on, as well as later, made it clear they know of these problems, the authors nevertheless plough ahead, ignoring these "irreducibilities" and do so with assertion rather than argument simply saying, for instance, that consciousness begins with a certain degree of brain complexity and did so many millions of years ago. Near the beginning they go into detail about features of consciousness in a section entitled "The four neuroontological subjective features of consciousness (NSFC)", which sounds impressive, but whose useful meaning entirely escapes me - are there any features of consciousness that are not subjective? And what does "neuroontological" mean? Slinging words together redundantly and creating jargonistic neologisms is no substitute for solid logic and argument, but it helps to convince an uncritical reader of the propriety of the endeavour, the labelling (NSFC), for instance, gives a sense that there really is some real phenomenon here to be studied scientifically.

The journey to complex organisms in evolutionary history is well described in great detail, with lots of tables and diagrams, but no evidence is adduced for the connection between any of this and consciousness, the word consciousness is simply slipped in. This not surprising since the authors did not define what they meant by consciousness or, crucially, how it could be publicly identified. Paragraph after paragraph juxtaposes mental events with objective data in plausible sounding sentences until one asks simple questions like, what on earth are they talking about? How do they know? Where is the publicly observable data?
The authors' assertion of consciousness simply piggy backs on the detailed and impressive account of evolutionary zoology and adds no new understanding, but "consciousness" gains a veneer of "scientificness " by appearing over and over again in the midst of all this impressive science. Dear reader, don't be taken in, this is page after page of sophisticated incoherence.

DUCK ERROR

One of their many mistakes is a familiar one to ethologists. They are not starting from the direct observation of publicly observable phenomena, but from an everyday concept, consciousness. As Bill Charlesworth would say, they are not following the Duck, but the theory of the Duck.

MIND AND BODY ARE NOT DIRECTLY CONNECTABLE

There are other errors which are worth discussing since they are far from unique to this book. Central to this is the relationship between, on the one hand, consciousness and subjectivity and mental events, and on the other physical phenomena such as observable behaviour and brain functioning. This relationship has taxed thinkers for centuries. Descartes (1649) tried to connect Mind and Body via the pineal gland. More recently there has been speculation that the connection between consciousness and brain physiology is mediated by "microtubles", or "quantum uncertainty" or "dendron" - "psychon" linkage via "quantum physics" (Lucas, 1961; Penrose, 1997 Eccles, 1992), or just by the fact we have a big brain (Greenfield and Collins, 2005). Feinberg and Mallatt bring in evolution.

By the late twentieth century, philosophical interest in study of consciousness had declined to a low level. But then in 1994, a young Australian philosopher, David Chalmers, electrified a conference in Arizona[2] by distinguishing between two "problems of consciousness", "Hard" and "Easy". What Chalmers called the "hard problem of consciousness" is this: how is it that activity of the nervous system can generate experiences - 'why should physical processing give rise to a rich inner life at all?' (Chalmers, 1995).

The "easy problems", which he believes will eventually be solved scientifically, include:

- 'the ability to discriminate, categorize, and react to environmental stimuli;
- the integration of information by a cognitive system;
- the reportability of mental states;
- the ability of a system to access its own internal states;
- the focus of attention;
- the deliberate control of behaviour;
- the difference between wakefulness and sleep.'

Whether the "easy problems", as expressed in the terms Chalmers uses, are indeed amenable to scientific solution is an open question, but that will not be addressed here.

The cleverness of Chalmer's distinction is that it again places (Hard Problem) consciousness, in the same category as the "easy" problems. The Hard Problem can thus be seen as just, well, harder, but not different in kind. The general notion seduces many, including Feinberg and Mallatt, into trying to describe an evolutionary progression in the same way as, for instance, the evolution of increasing brain size might be described. Why is this an error?

The short answer is that subjective experience has, for more than a century, been seen as not something amenable to a scientific approach. The behaviourist J.B. Watson unequivocally rejected the use of subjective findings in science:

> "One can assume either the presence or the absence of consciousness anywhere in the phylogenetic scale without affecting the problems of behaviour by one jot or one tittle; and without influencing in any way the mode of experimental attack upon them". (Watson, 1913).

[2] The somewhat oxymoronically called "Association for the Scientific Study of Consciousness" was formed after this meeting.

Tinbergen (1951) also rejected the use of "subjective phenomena".

> "Because subjective phenomena cannot be observed objectively in animals, it is idle either claim or to deny their existence." (page 4)

Tinbergen also points to the inappropriateness of emotional explanations, for instance, to say that an animal is "caused" by hunger to seek food and eat. This confuses a psychological with physiological explanation; hunger, he says, is just a "convenient description" (page 4). To unpack the point further: the description is circular:

Why does the animal look for food?	Because it is hungry.
How does one know the animal is hungry?	Because it looks for food.

Both authors do not so much deny the existence of subjective phenomena, as point to the futility of trying to make them the object of scientific study.

ASCRIPTION NOT DESCRIPTION

Watson's choice of words is interesting, "One can **assume** either the presence or the absence of consciousness". He is right on two counts. Firstly consciousness is not **de**scribed it is **a**scribed, he used the word "assume", which in this context means the same as ascribe, both refer to the attitude or stance of the ascriber/assumer not the object of their assumption. Secondly, the assumptions / ascriptions can be made without affecting the scientific understanding of the behaviour. They are irrelevant. There is one caveat to this second point. Ethologists frequently ascribe feelings to animals, but this is not as part of their data gathering but as a heuristic helping them to decide what to measure. Tinbergen, by saying it was "idle" to claim or deny subjective phenomena, was implying that such claims are in a different realm of discourse.

OTHER MINDS

The longer answer as to why Feinberg and Mallatt, like so many others, make this error will involve not only showing why it is misguided but also it will involve offering alternative approaches to the study of consciousness. As Kuhn (1962) pointed out, a scientific paradigm does not fall simply due to contradictory evidence or even some internal incoherence, it needs an alternative paradigm to take its place. I have discussed possible alternative approaches in an earlier paper (Richer, 2016) and so will not present the whole argument here.

Let me start with why the "scientific study of consciousness" in the sense Feinberg and Mallatt intend it, is an oxymoron. Its central problem lies in how we arrive at some piece of knowledge.

So how do I know that other people have a first person perspective, that they have experience? I don't, I assume it, I ascribe it to them, as I have already said. How do I know other species have experience. Ditto, I don't. But to some I ascribe it. The thought is hardly novel, it is the centuries old philosophical problem concerning the existence of "Other Minds". Husserl (e.g.1913) was exercised by the essential difference between one's own experience and that of others. Slightly more recently, Freud, for example, wrote in 1915.

> 'Consciousness makes each of us aware only of his own states of mind. That other people, too, possess a consciousness is an inference which we draw by analogy from their observable utterances and actions, in order to make this behavior of theirs intelligible to us.' (Freud, 1915).

Many philosophers have discussed the difference between talking about the mental states of others compared to talking about (objective) natural phenomena. Wittgenstein (1953) wrote that talking about someone's mental states was to take a stance rather than discriminating logically or factually. Strawson (1962) contrasted "ordinary interpersonal attitudes" with a "purely objective view", i.e. the issue is *approached* in two different ways. So we treat biological machines, a.k.a. other people, as conscious beings, simply because we have (unconsciously) decided to.

AGREEMENTS

But there is a stronger argument. It rests on the necessary priority of agreements over knowledge. It is this. Any corpus of *shared* knowledge consists of *agreements* between individuals. If there are no agreements there is no *shared* knowledge. Private knowledge does not require that sharing, but then it cannot be part of the shared body of knowledge, and so is not something, *a priori*, that <u>we</u> can talk about and be understood.

Just as fish are, as the saying goes, the last to discover water, this simple point has a significance which seems to elude many. Issues around consciousness become clearer when we realise that <u>we</u> must give intellectual priority to <u>agreements</u>, which can only be public, and not focus on <u>knowledge</u>, which can be both public and private.

Two types of agreements may be distinguished: Negotiated (N type) and Demonstrated (D type) (Richer, 1975). An example of Negotiated agreements is: we agree that this colour means "red", or that "flower" means this sort of thing and we agree on the use of "this" and "is". What words (in the sense of the sounds we make or the marks on paper) are used is largely arbitrary, (and different in different languages) and the exact boundaries of the meanings are sometimes fuzzy and may be different in different languages (Nagel, 2014).

Possessing these Negotiated agreements enables us to proceed to Demonstrated agreements. A Demonstrated agreement is of the type, for example, "this flower is red". Having agreed the meanings of each word we can agree, or disagree, on "this flower is red" simply by demonstration, perhaps by pointing at it.
Negotiated agreements can be changed and are judged by whether they are *useful*. Demonstrated agreements are judged by whether they describe the real shared world accurately, by whether they are, in this sense, *true*.

Type or agreement	*Evaluated according to*
N Negotiated	Usefulness
D Demonstrated	Truth

This is relevant to consciousness in the following way. Thinkers have become side tracked into asking questions about conscious knowledge, and easily thus confuse shared (public) knowledge with private knowledge (the private conscious experience), which cannot be discussed until it is shared.

AGREEMENTS IN SCIENCE

This agreement process is formalised and tightened in science: terms are carefully and precisely defined, or at least the attempt is made to do so, and the phenomena under study are described in these terms. A hypothesis couched in these terms can be tested by an attempt at demonstration, such that, having accepted the terms, the scientists can agree by demonstration on the result. Science, as Medawar (1967) argued, is about creating a coherent story about the facts and which can be tested by seeing whether it accords with the facts. The facts are the demonstrated agreements.

So when someone asserts that some organism is conscious, or when they talk about what someone is conscious of, the individual experiences, where and how do these fit into science? The answer is, they don't. The assertion that some organism is conscious, or the nature of contents of consciousness, cannot be agreed by demonstration. The viewpoint of one of the people (or other animal) involved, is, firstly, qualitatively different from that of everyone else (as Feinberg and Mallatt discuss) and secondly, without that one person, the phenomena in question (the conscious experience) would not exist; as when someone is dead. Conscious experience cannot be agreed by demonstration in the same way that, say, "This flower is red" can be agreed.

ONTOGENY OF THE USE MENTAL STATE WORDS

So how are mental state words agreed? This can be rephrased as: how do children learn to use these words? Learning how appropriately to ascribe a feeling or other mental states to someone else is the same as learning the meaning of the word "flower" or "red", save for one vital additional aspect. The similarities are that child comes to recognise, after observing others (being an onlooker), that a certain matrix of situations and behaviour is labelled by certain words. He learns the "public criteria" (Wittgenstein, 1953) for mental state words. If we could not sufficiently identify a state by public phenomena (behaviour and context), there would be no shared meaning. So two people, being in the position of onlookers, can say of a third, "he is excited", and the word "excited" can have some shared, albeit imprecise, meaning.

But when someone says, "I am excited" they do not look in the mirror, they feel it, they experience it. Here their viewpoint is that of the agent, the person who takes action. Feedback from their motor output and other reactions may play a part in that experience. But their viewpoint remains uniquely that of the agent. As well as learning the public criteria, child comes to recognise the feeling states so they can they apply a mental state word to themselves. They can say, "I am …". The agent and onlooker viewpoints are intimately intertwined within the meaning of single mental state word – "I am furious" is intertwined with "you are furious" and "he / she is furious".

Children start to interweave agent ("I") and onlooker ("you". "s/he") viewpoints from a very early age, and this is part of the development of their intersubjectivity/mentalising skills. These mentalizing skills are developed in rudimentary form by about 18 months, but continue to become more sophisticated and fluently used throughout childhood. The child negotiates shared meanings (Newson & Newson,1975) and is drawn into the shared mind which is their culture. In particular, they negotiate the meanings of words describing mental states. An exception is when these social interactions are hardly practiced. This is the case with children who behave autistically. Indeed such children throw into the sharp relief the integration of first-person and third/second-person perspectives, the integration of Self and Other, precisely by their difficulty in doing just that (Williams, 1996; Richer, 1978, 2001).

In our daily use of mental state words and our interweaving of the viewpoint of Self and Other, we are so fluent that it is usually forgotten that the meaning of these words is only negotiated, not demonstrated. They are essential social tools which bind us together and enable sophisticated communication. But part of their meaning is not publicly defined or identifiable in the way that something agreed by demonstration is. They are not data or facts in the way that publicly observable behaviour is.

AGENT'S AND ONLOOKER'S STORIES ABOUT HUMAN BEHAVIOUR

So a distinction must be made between two types of research

(i) studying humans including the assertion of consciousness and the contents of consciousness, their experiences. This is the story which embraces agency.

(ii) studying observable human behaviour including the fact that humans have the concept of consciousness and share first-person perspectives. This is the onlooker's story, and it is what science generates.

The first type of research would include studying the contents of consciousness, i.e. feelings, thoughts, intentions, etc., the first-person perspective. This has been firmly rejected as not possible or useful as part of a scientific approach. That does not imply such study is worthless, far from it. It gives our lives meaning, colour, beauty, direction and connection with others. It has evolved and has great adaptive value. It is what we do naturally.

The second type of research includes studying phenomena around the observation that humans have the concept of consciousness and, in particular and crucially, seem to be able share thoughts and feelings with each other, albeit often miscommunicating. This phenomenon has various names: Intersubjectivity, Mentalization, Mindreading, having a Theory of Mind, an ability to Emphasise. I prefer the terms Intersubjectivity (e.g. Trevarthen & Aitken, 2001, 2003) and Mentalization (e.g. Fonagy et al, 2011); intersubjectivity since it captures the sharing (inter) of subjective phenomena (subjectivity), and Mentalization since that embraces the way we treat each other as agents like ourselves and how failure to do that is associated with numerous mental disorders and other dysfunctions. The evolution and ontogeny of this ability has been discussed elsewhere (Richer, 2016). Pursuing this approach is difficult and require considerable intellectual discipline since it is so easy to slip into the logic of the agent's story, it being how we have evolved to think.

The idea that there two separate stories to be told about human behaviour is far from new. Brentano (1874) distinguished "genetic psychology" which uses third-person only, from "descriptive psychology" which also embraces first -person narratives. Dennett (1987) contrasted physical and teleological stances on the one hand with an intentional stance on the other. The anthropologist and linguist Pike (1967) made a related point when he distinguished an "emic" account, which uses the culture's own terms and so will include first-person narratives, from an "etic" description of a culture, which uses the anthropologist's own terms. Some mainstream psychology has got itself into difficulties by muddling these two approaches (e.g. Richer, 1975, 2020). Feinberg and Mallatt get into similar muddles.

REFERENCES

Brentano, F. (1874). *Vom Ursprung sittlicher Erkenntnis (The Origin of our Knowledge of Right and Wrong)*. Meiner, Hamburg. DOI

Chalmers, D. (1995). Facing Up to the Problem of Consciousness. *Journal of Consciousness Studies, 2*(3), 200–219.

Dennett, D., C. (1987). Three Kinds of Intentional Psychology. In: Dennett, D. C., *The Intentional Stance*, pp. 43-68. The MIT Press, Cambridge.

Descartes, R. (1989). *Passions of the Soul.* Indianapolis: Hackett.

Eccles, J.C. (1992). Evolution of Consciousness. *Proceedings of the National Academy of Science, 89* (16), 7320-7324. DOI

Fonagy, P., Bateman, A. & Bateman A. (2011). The Widening Scope of Mentalization: A discussion. *Psychology and Psychology and Psychotherapy: Theory, Research and Practice, 84*(1), 98–110. DOI

Freud, S. (1989). *The Freud Reader (P. Gay, Ed.)*, pp. 575. WW Norton & Co, New York.

Globus, G. (1973). Unexpected symmetries in the "world knot". *Science, 180(4091),* 1129-1136. DOI

Greenfield, S. A. & Collins, T.F. (2005). A neuroscientific approach to consciousness. *Progress in Brain Research, 150,* 11-23. DOI

Husserl, E. (1913). *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie.* De Gruyter. DOI

Kuhn, T. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lucas, J. R. (1961). Minds, Machines and Gödel. *Philosophy 36*(137), 112-127. DOI

Medawar, P.B. (1967). *Science, the Art of the Soluble.* Methuen, London.

Newson, J., & Newson, E. (1975). Intersubjectivity and the transmission of culture; on the social origins of symbolic functioning. *Bulletin of the British Psychological Society, 28*, 437-446.

Penrose, R. (1997). On understanding understanding. *International Studies in the Philosophy of Science, 11*(1), 7-20. DOI

Pike, K.L. (1967). *Language in Relation to a Unified Theory of the Structure of Human Behavior*. Mouton, Elmsford, New York: Glendale Summer Institute of Linguistics. DOI

Richer, J.M. (1975). Two types of agreement - two types of psychology. *Bulletin of the British Psychological Society 28*. 342-345. (Reprinted in Readings in psychology for the Teacher Child D. (ed) Holt, Rinehart, & Winston, London 1977).

Richer, J.M. (1978). The partial non communication of culture to autistic children. In: Rutter, M. and Schopler, E. (eds) *Autism: reappraisal of concepts and treatment.* Plenum Press. New York. (paper previously presented in 1976 at autism conference, St. Gallen, Switzerland)

Richer, J.M. (2001). The Insufficient Integration of Self and Other in Autism. In: Richer, J.M. and Coates, S. (eds) (2001). *Autism: The Search for Coherence*. London; Jessica Kingsley.

Richer, J.M. (2016). Mentalistic and scientific stories about human behavior, biomimetic heuristics and psychology's confusions. *Human Ethology Bulletin, 31*(4), 15-33. DOI

Richer, J. (2020). Comparisons between Mainstream Social Psychology, "Examples" Psychology and Ethology stimulated by Michael Billig's book: More Examples, Less Theory. *Human Ethology, 35*, 137-156. DOI

Strawson, P.F. (1960). Freedom and Resentment. *Proceedings of the British Academy, 48*.

Trevarthen C, & Aitken KJ. (2001). Infant intersubjectivity: research, theory, and clinical applications. *Journal of Child Psychology and Psychiatry, 42*(1), 3-48. DOI

Watson, J.B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*, 158-177.

Williams, D. (1996). *Autism: an Inside-Out Approach*. London, Doubleday.

Wittgenstein, L. (1953). *Philosophical Investigations.* Oxford, Blackwell.