Review of: "Growing Confidence and Remaining Uncertainty About Animal Consciousness"

Harry Smit

Potential competing interests: No potential competing interests to declare.

Irwin's article is an attempt to review the many ideas that have been discussed within the field of consciousness studies. The core problem discussed in this field is how we can bridge the gap between brain processes and conscious, subjective experiences. Irwin argues that there is growing empirical evidence that consciousness requires a complex brain, but also that there are uncertainties. These uncertainties are related to the conceptual problem of how subjective experiences are explicable in terms of the brain. An asset of Irwin's review is that it incorporates a discussion of the results of comparative studies of the nervous systems. These studies may help us in answering the question of when consciousness evolved in animals. Although there are no clear answers, Irwin is optimistic: he expects that the question of how consciousness arises (and arose during the course of evolution) in more complex nervous systems will be answered if a more suitable vocabulary is developed. The latter is a requisite for resolving the conceptual problems.

In the following, I will briefly discuss an alternative conception not discussed by Irwin, namely the neo-Aristotelian conception extended with evolutionary theory, as a possible explanation of (the evolution of) consciousness. I assume that Irwin has good reasons for ignoring this alternative conception, but I hope to convince him that, by comparing it to his ideas, it makes sense to reconsider his preferred (evolutionary) explanations.

The neo-Aristotelian conception is advanced as an alternative to the neo-Cartesian conception of consciousness. The latter has been the dominant conception since the scientific revolution. The reason is that the new physics proved the Aristotelian teleonomic conception of physics to be explanatorily unfruitful. Efficient causal explanations (not in teleonomic but in nomic terms) were shown to be adequate for physics. Descartes later added that they should be adequate for life science as well. Natural phenomena other than human thought and consciousness should, according to him, be explained in terms of the mechanical laws of matter in motion. By excluding consciousness from the study of nature, Descartes shaped a new conception of the mind. Having a mind is, according to Descartes, to have experience and to be aware of oneself as having experience. He argued that animals do not have experiences and are not self-conscious: they are to be explained in terms of the mechanical laws of matter in motion.

What were the arguments of Descartes for redefining the mind in terms of consciousness? The starting point in his reasoning is that he (and his predecessor Galileo) noted that the results of the new physics are at odds with mundane experience. Although the world appears to be, for example, multi-colored, in reality, there are only rapid-moving, colorless particles, waves, or electromagnetic radiation. Qualities like color do not in reality belong to objects, for they are, according to Galileo and Descartes, purely the offspring of the mind. They belong to subjective experiences. Hence, we

find in the writings of Galileo and Descartes the origin of the gap between the objective, physical realm and the subjective mental realm (the world of conscious experiences). This gap is still haunting scientists today, as Irwin's review shows.

Neo-Aristotelians have explained that Galileo's and Descartes' arguments are for many reasons defective. I mention only two of them. First, Galileo and Descartes argued that only numerical measurement is a mark of objectivity (this argument can be traced back to Plato and Pythagoras). We can, for example, measure the length and weight of objects, but not their colors (they are therefore not the qualities of objects in reality). Consequently, the domain of science was delimited by Galileo and Descartes to investigations of space-occupants. However, although colors are not susceptible to numerical measurement, it does not follow that they belong to the subjective, mental domain. After all, we use color samples (standards of comparison) for determining the color of objects. The way parents teach their children the role of standards for the correct application of color words is illustrative. For example, caregivers may explain the meaning of the word 'red' by referring to a ripe tomato. If a parent points at the tomato and utters the sentence 'This is red', then he or she gives a so-called ostensive definition (a rule). It can be paraphrased as 'This color is red' or 'This color is called red'. The tomato is then used as a sample (a standard of comparison belonging to the means or method of representation) for explaining what red is and can be used to determine whether other objects in the surroundings are red too. They are correctly said to be red if they are the color pointed at. The functioning of a color sample as a standard is comparable to the role of the meter stick as a sample. And just as a meter stick functions as an instrument belonging to the method of measurement, so too the tomato functions as an instrument belonging to what measures, not to what is measured. The point to notice is that the meaning of 'red' does not originate in having subjective sensory experiences (as Cartesians assume). There is no such thing as explaining a color by referring to inner, subjective experiences.

Second, Galileo and Descartes argued that colors belong to the subjective, mental domain because they, just like sensations like pain (or tactile perceptions), are *felt.* Yet the problem with the use of this analogy is that, when we perceive, for instance, a red rose in the garden, we do not have any sensation somewhere in our mind or brain (comparable to the sensation of, for example, a headache). We see the color of an object in the external world, not in our mind, and we do not feel anything (while we feel a pain). To have a visual perception of a red rose is simply to see it.

This criticism of the ideas of Galileo, Descartes, and their current followers has some interesting consequences. If it is correct to say that Galileo and Descartes took a cul-de-sac, then we can ask what the correct alternative is. The perhaps difficult point to notice here is that what we have to explain is not how subjective, conscious experiences are generated in the brain (as Irwin also assumes), but how our ancestors started to use samples and color words and whether we can explain (in terms of modern evolutionary theory) how this normative practice emerged. Hence, there is a change in the explanandum. The alternative explanation has already been discussed by neo-Aristotelians: our ancestors started to use color samples and the associated perceptual vocabulary as an *extension* of their natural (innate) discriminatory abilities. The example of red-green colorblind persons can be used as clarification. They cannot, as the result of mutation, discriminate between green and red, but they also cannot use our rules (involving the use of samples) for the use of the words red and green and the associated perceptual vocabulary. That is unsurprising, since children do not only respond to what they see but also learn to express what they perceive with the aid of a perceptual vocabulary.

I have discussed perception, but similar arguments can be given for other mental phenomena. I mention emotions. Irwin asks how and when subjective emotional feeling or affective experience arose from the material substrate of a nervous system. The focus is again on subjective experiences: when could organisms distinguish external stimuli from internally generated subjective feelings? Neo-Aristotelians, by contrast, follow Darwin (1872), who noted that we can investigate the evolution of the expressions of emotions by studying the different forms they take. For example, teeth-baring in dogs evolved as an expression of threat when a receiver noticed that the baring of teeth preceded what the dog was going to do next (it was a cue for an attack). By noticing the bared teeth, the receiver could anticipate the attack and could, for example, flee or start threatening before being attacked (reducing the fitness costs of injuries). Thus, as the result of selection on both sender and receiver, teeth-baring evolved into a separate threat signal. We can understand the evolution of smiling in a similar way. It is nowadays, just as laughter, often an expression of amusement and enjoyment, but is presumably derived from an expression of submissive behavior (itself derived from a fear response) displayed by our ancestors.

That emotions are manifest in facial expressions clarifies why we hesitate to ascribe emotions to animals such as fish, reptiles, and amphibians, for their behavioral repertoire does not include expressive behavior. They lack a rich facial musculature: we do not see facial expressions of anger or fear that warrant the ascription of an emotional response when they attack or flee. By contrast, the snarl of a dog exposing its fangs coupled with a stiffened tail and ears drawn back warrants ascription of an emotional response to the dog when it is displayed on an appropriate occasion.

It is of course not denied that emotions, like sensations, are feelings. Yet neo-Aristotelians and neo-Cartesians conceive of feelings in a different way. Feelings are, according to the Cartesian conception, not behavior but something subjective underlying behavior. For Descartes distinguished an inner reality with which we are intimately acquainted from outer bodily movements. Consequently, if we want to understand and explain emotions, such as fear, joy, or affection, we should, according to Descartes, focus on this inner mental reality. Yet there is a price to pay: the evolutionary origin of these subjective experiences becomes then deeply puzzling: how can we explain their origin if they are categorially distinct from the transactions occurring in the brain? The difference between the neo-Cartesian and neo-Aristotelian conception is that, while the Cartesian conception emphasizes that having an emotion signifies experiences *of which* we are conscious, neo-Aristotelians emphasize that having an emotion signifies that we have or enjoy experiences and that emotions are mental states we are in *while conscious* (in opposition to when we are asleep or unconscious). According to neo-Aristotelians, it is misguided to argue that we are conscious of an emotion since there are no inner sense organs by means of which we become aware of our emotion or emotional feeling. This difference, which perhaps appears to be a minor one, is a major one. Let us explain.

Neo-Aristotelians, in contrast to what neo-Cartesians suppose, argue that what we see of an emotion is not its external signs. Smoke may be conceived as the external sign of fire, but if someone is shouting, gesticulating, trembling, raising a fist, or banging the table, it would be mistaken to conceive them to be the mere external signs of anger. Although there is a difference between an expression of an emotion and the emotion itself, there is an internal relation between the form of an emotion and what it is a form of. Shouting and gesticulating are (part of) the forms the emotion takes (but not

necessarily always takes). They are not the effects of someone's anger (while smoke is the effect of fire). Hence, it is a misconception of neo-Cartesians to argue that the relation between behavioral manifestation (form) and the emotion is an external relation, implying that the emotion is hidden and an inner cause of exhibited behavior, as if the behavior is merely an external sign of what is concealed from sight. The emotion is not hidden but visible in someone's behavior. The behavioral and, when children have mastered the use of language, linguistic manifestations of emotions are *criteria* for ascribing emotions to human beings.

I said earlier that the neo-Aristotelian criticism of the neo-Cartesian conception leads to a change in what has to be explained. What we have to explain is not how we can understand subjective experiences or feelings as products of a complex nervous system, but how intransitive consciousness, different forms of transitive consciousness, and self-consciousness (in humans alone) evolved. Explanations have been discussed elsewhere (for example, Bennett and Hacker, 2022, chapters 10-14; Smit and Hacker, 2020) and will not be repeated here.

I have discussed in this comment an important distinction: as long as we can distinguish between a specification of a measure (a sample like a ripe tomato functioning as a standard of comparison for determining whether objects are red, a specimen for determining whether an individual belongs to a certain species, or a drawing or picture of a basic emotion functioning as a standard of comparison) and the statement of a measurement, we can distinguish what is a priori (or conceptual) and what is empirical. That distinction should, according to neo-Aristotelians, replace the exceedingly opaque distinction between objective and subjective (two terms that should be eschewed). If we take this neo-Aristotelian turn and not the neo-Cartesian cul-de-sac, then we can without any ado construct evolutionary explanations of how consciousness evolved. At the theoretical level, there are no changes. Yet, as we have seen, the change in the explanandum involves a conceptual revolution: we need to abandon the neo-Cartesian conception.

References:

Bennett, M.R. and Hacker, P.M.S. (2022) *Philosophical foundations of neuroscience*, 2nd revised edition. Blackwell Publishing, Oxford.

Darwin. C. (1965 [1872]) *The expression of emotions in man and animals* Chicago University Press, Chicago.

Smit, H. and Hacker, P.M.S. (2020) Two conceptions of consciousness and why only the neo-Aristotelian one enables us to construct evolutionary explanations. *Humanities and Social Sciences Communications*. 7:93. https://doi.org/10.1057/s41599-020-00591-y.