E

Evolution and Self-awareness

R. Haven Wiley
Department of Biology, University of North Carolina, Chapel Hill, NC, USA

Synonyms

Consciousness; Self-consciousness

Definition

An ability to think about one's own mental states.

Introduction

It seems clear to me that my subjective mental states are my own. Furthermore, this act of introspection shows that I have some ability to think about my mental states. I would say that I am self-aware. Furthermore, my experience indicates that most humans have such mental states of their own, including self-awareness. The question thus arises where do these subjective mental states, including self-awareness, come from? What causes or explains their presence and content? Has this human capability evolved?

From the earliest times, discussions of consciousness have been closely related to those of free will. Volition is still for many people a definitive attribute of consciousness, which in turn is a definitive attribute of humans. Both volition and consciousness, in turn, are closely allied with rational thought. Often these three capabilities are combined as the distinctive characteristics of the human soul. For many people, the attributes of these capabilities seem entirely apparent by introspection. For all of these reasons, discussions of self-awareness and volition have often become doctrinaire, for instance, in both Judeo-Christian-Islamic-Vedic morality and Brahmanic-Buddhist-Taoist-Gnostic transcendence. Human and nonhuman animals have seemed to occupy their respective rungs on a *scala naturae*, each of which had its distinctive attributes augmenting the rung below and subsumed by the rung above. Consciousness and its concomitant volition and rationality were the attributes that separated humans from all lower forms of life.

Modern western philosophical and theological discussions began in the sixteenth and seventeenth centuries with fundamental controversies between Christian sects about divine grace and predestination and with Descartes' subsequent dictum "Je pense donc je suis." In recent centuries, contrasting positions have developed between environmentalist and nativist theories of human perception (or between sensory and mental determinants of thought). These trends have culminated in an emphasis either on culture or on intrinsic structure as the predominant determinants of consciousness and language. In all of

T. K. Shackelford, V. A. Weekes-Shackelford (eds.), *Encyclopedia of Evolutionary Psychological Science*,
https://doi.org/10.1007/978-3-319-16999-6_2202-1

this history, because it was assumed that the uniqueness of humans depended on consciousness and self-awareness, there was little attention to the possibility of their evolution.

Continuity Between Human and Nonhuman Animals

Darwin's revolutionary book, *The Expression of the Emotions in Man and Animals* (1872), for the first time, proposed some continuity between the minds of humans and other animals as indicated by their respective behaviors. Since then, studies of learning and instinct, as well as the ontogeny and phylogeny, of behavior have had progressively increasing influence on thinking about consciousness and awareness. Particularly relevant have been experiments that explore the limits of nonhuman intelligence. These experiments have raised difficult questions about the relationships among language and thought and learning and consciousness.

For instance, studies of a number of nonhuman animals (especially chimpanzees, other apes, parrots, and dolphins) have indicated that these organisms can respond to complex stimulation, such as encoded queries and requests, in ways that resemble our own use of language (Savage-Rumbaugh et al. 1998). Just as we routinely attribute consciousness and will to other people by empathy based on their behavior, we are inclined to attribute these capabilities to the subjects of these experiments. These studies are unusual among biological and psychological experiments in two ways: their small samples of subjects and their intensive involvement of humans (as opposed to experiments with many comparable subjects and largely mechanical interactions with them). The small sample is justified because the interest is in the potential rather than the norm of behavior. If one individual chimpanzee can converse like a human, that alone makes the point.

The lack of replication in combination with intensive human interactions with subjects raises further issues. Each of these studies has been conducted in one laboratory and directed by one principal investigator or a small team. No team has ever carefully replicated the procedures of any other. Nevertheless, there has been an accumulation of similar results with similar protocols, so that it is becoming more difficult to exercise broad skepticism about the kinds of responses evoked.

There is an even deeper controversy about language-like responses by nonhuman animals. Interpretations of these responses often devolve into a polarity between attributing the observed responses to no more than thorough (rote) learning as opposed to the spontaneity or creativity that human language seems to show (Lutz 2009). There are twin problems here – it is difficult to distinguish complicated from random patterns of responses, and it is also hard to distinguish repeated from rote responses. Infrequent, unprecedented, and unrepeated responses, just what we look for in consciousness, cannot be easily shown statistically to differ from accidental or random responses. On the other hand, statistically significant patterns of response require replicated results, which then come to resemble thoroughly learned or rote responses.

The issue of thorough learning, as opposed to volitional thought, is a pervasive problem in comparative studies of consciousness. Consider another example of convergence in mental abilities of human and nonhuman animals – abilities to respond to oneself in a mirror and to attribute mental states to others (and by extension therefore to oneself). Responses to mirrors are no doubt tricky. Many animals respond to mirror images (itself a remarkable capability) as if the image were another individual of its species, perhaps a rival evoking aggression. Chimpanzees in contrast behave as if they see themselves in a mirror, for instance, by touching unexpected marks on their faces that they see only in a mirror. Such responses to mirrors, which seem normal to most humans nowadays, indicate a remarkable advance in mentality. Yet this ability clearly requires learning. Humans with no experience of mirrors or photographs of themselves, as anthropologists often report, do not easily recognize their images. And anyone can try the experiment of directing movements (or even identifying oneself) in a *mirror image of a mirror image* of yourself, an experience that reveals uncomfortably that mirrors require considerable practice to master. Even recent experiments in which chimpanzees without a direct line of sight can use a mirror to sign to a human recipient, but do so only when the recipient is looking at the mirror (Lurz et al. 2018), are subject to the same questions about thorough learning. Evidently humans nowadays master some superordinate associations involving mirrors that most nonhumans have not, but humans have not mastered all the possible associations. Nor is it easy to determine whether or not some nonhuman animals have comparable mastery (De Veer and van den Bos 1999).

Actions that result from rote learning raise another question about consciousness. Humans often master well-practiced tasks to the extent that they are performed unconsciously. Conspicuous examples are actions or sequences of actions performed routinely, such as making a cup of coffee each morning. Yet on any one occasion, the awareness of having performed an action can escape us, for instance, when we are perplexed about whether we had already added the sugar. Certain actions, swallowing and walking, for instance, consist of complex muscular coordinations that we are seldom aware of, although this lack of awareness can cause serious accidents.

Furthermore, this issue of the criteria of consciousness merges with the issue of private sensations. How can we know, other than by imputations based on coarse empathy, what another organism feels or even senses? Introspection is the source of these insights. How can it be determined that another organism, other than myself, is conscious, acting by will rather than rote, in the same way I do?

Neurophysiology of Consciousness

An obvious possibility for recognizing consciousness is to investigate neural activity during presentations of stimulation. Neurobiologists can now detect in detail the neural events that result from sensory stimulation, both those sorts that we normally are aware of, for instance most exteroceptive sensation, and those we are normally not aware of, for instance most proprioceptive sensation. This approach can extend to neural events in the brain. Physiologists now know a great deal about local areas in the brain, even particular neurons, that are specialized for analyzing sensations, controlling muscles, generating emotions, consolidating memories, comprehending or producing language, even recognizing a visual pattern as a human face. Centers in the brains of nonhuman animals have similar functions (the centers for learning and producing songs by birds are particularly well documented). Is there such a center for consciousness?

Suggestive in this case is an experiment that seems to reveal a half-second or so delay between the initiation of a spontaneous action, on one hand, and awareness of it, on the other (Libet et al. 1983). Because action precedes awareness, it appears that action triggers awareness, rather than vice versa, so that consciousness is the effect of our actions rather than their cause. Volition becomes an illusion, and consciousness seems to be a distinct operation, a candidate for localization in the brain. Some qualifications are in order, though. Subjects record their awareness of the action by remembering the exact position of a spot moving rapidly around the face of a clock. Recording this visual stimulus in memory is itself a response to the spontaneous decision to act, just as much the act itself is a response to this decision. Such an experiment thus does not necessarily reveal that action precedes volition. Instead each of these two operations requires different neural events lasting finite, and evidently not exactly equal, amounts of time. Furthermore, the memory is encoded in language, which becomes the sole means of obtaining the datum actually recorded by the experimenter. Again we are back to questioning how we can know what another organism feels or thinks, unless that organism tells us in some way.

The inability so far to find a locus in the human brain specialized for consciousness has led to proposals that awareness results from distributed networks of neural interactions. Computer programs for learning complicated tasks often employ “neural networks,” one or more intermediate layers of “cells” that reciprocally influence each other’s activity, between an input (sensory) layer that provides initial conditions and an output (motor) layer that represents the response. Further programming defines the utility of any response and uses this evaluation to regulate the stability or variability in the properties of cells in the intermediate layers. An arrangement that produces responses with low utility is thus allowed to change (mutate) before subsequent trials; those that lead to responses of high utility are stabilized (saved), in other words, learned. Although these computer programs are called "neural networks," it is still not clear how closely they resemble operations in a brain. Only at a superficially general level can we suppose that distributed operations in the brain share the features of computational "neural networks." The specifics of neural processing to produce consciousness remain as elusive as ever.

Continuity Between Brains and Other Machines

The relationship between consciousness and language arises in proposals to distinguish humans from other machines – or by extension to determine whether or not any machine is conscious. Turing's test and Searle's modification of it are examples (Dennett 1991; Searle 1997). Each involves a judge posing problems to unseen contestants. The issue is whether a human (conscious) contestant can be distinguished reliably from a nonhuman (unconscious) one. Searle contends that Turing's test would not distinguish between a human who understood a language and one who just followed rules by rote. It thus could not distinguish a conscious human from an unconscious machine. By extension, it is worth emphasizing, it would also not distinguish between a conscious and an unconscious machine. A fundamental question here is whether or not conscious behavior, such as language, is strictly rule-following or not. And thus whether or not humans are strictly rule-following machines or not. We might also extend this question to whether or not machines do or do not strictly follow rules.

Noise as a Determinant of Consciousness

These conundrums about the relationships of volition, consciousness, and learned and unlearned behavior all intertwine with issues of language and even communication in general. Any evidence about an organism's consciousness depends largely, in the final analysis perhaps exclusively, on what it reveals in its behavior. The evolution of consciousness thus depends in a fundamental way on the evolution of communication. It is thus remarkable that noise influences the evolution of communication in a way that provides a straightforward explanation for the evolution of subjective experience (Wiley 2015).

Noise, as Claude Shannon first emphasized in his revolutionary analysis of information, is anything that results in errors by receivers during communication (Shannon and Weaver 1963). Noise can consist of extraneous irrelevant background stimulation that mixes with signals during transmission from signaler to receiver. This is what is commonly thought of as noise. Noise, in Shannon's sense of errors by receivers, can also result from degradation and attenuation of signals during transmission. It can also result from irregularities in a signaler's nervous system which introduce irregularities in its signals; and it can consist of analogous irregularities in a receiver's nervous system which introduce irregularities in its perceptions. As a result of any of these sources of noise, a receiver's perception only imperfectly reflects a signaler's actual situation.

To apply this approach to the evolution of consciousness, note that the dilemma confronting a receiver of signals in communication is strictly analogous to that confronting a perceiver of external objects and events in general. Noise in perception can result from mixing of sensations from irrelevant sources, from degradation and attenuation of stimulation during transmission, and from unpredictability in a perceiver's own nervous system.

At the moment of perception, a perceiver has no way to determine whether or not the perception corresponds to a particular external situation or to an erroneous illusion. All the perceiver knows at the moment is its perception. Nevertheless, memory of repeated perceptions, especially in combination with communication with other individuals, could reveal these discrepancies. In this way such an organism, capable of thought and language, could develop a sense that its own perceptions differed, in some respects and on some occasions, from those of others. Both some ability for abstract thought (a capability for generalization and discrimination) and some ability for communication of such abstractions seem crucial for this awareness. Otherwise individuals would be isolated within the shell of their own perceptions. They might well learn to avoid or to prefer certain perceptions, but it would be difficult to compare them with other individuals'.

Because of noise in perception or communication, a perceiver or receiver must make a decision every time it acts on any sensation. It must decide whether the sensation warrants a response (and also which response). In other words, it must decide whether a sensation is a signal (with some relevance for the perceiver) or noise (with no, or misleading, relevance). Noise creates the unavoidable possibility of two incompatible kinds of errors in perception, false alarm or missed detection. All perceivers, even those organisms such as sponges or bacteria with no nervous system like ours, are perforce decision-makers.

Signal detection theory (Macmillan and Creelman 1991), based on Shannon's theory of information, and decision theory allow a formal mathematical analysis of the performance of any perceiver in the presence of noise. Because of the two conflicting sources of error, a perceiver is in a double bind. It cannot reduce one source of error without increasing the other. As a consequence, it can only optimize its decision in particular circumstances and cannot attain perfect performance. This optimization leads to a fundamental conclusion that *perceivers cannot escape from noise*; their only option is to optimize their responses in each situation.

This is what organisms actually do. All must deal with unpredictable contingencies. Evolution by natural selection provides a mechanism that can optimize, within limits, neural capabilities to make decisions that promote survival and reproduction for the organism. In a fundamental way, nervous systems are decision-making organs devoted to this task of responding efficiently to conflicting possibilities of stimulation. Every organism must confront its subjectivity with some decisions, no matter how crude the mechanism.

Awareness of subjectivity in perception, however, requires a nervous system to form higher-order associations. When these connections between subjectivity and objectivity reach awareness, we can expect consciousness. The process requires a sufficiently complex nervous system. The logical inconsistency of self-reference might indicate that no such system can ever be completely aware of all its operations, even its degree of self-awareness.

Humans clearly have achieved the highest performance so far. Yet the evolution of this capability seems likely to have emerged gradually by successively more complex mental associations. Whether or not other organisms (great apes and bottle-nosed dolphins come to mind as possibilities) have reached states of consciousness comparable (although perhaps not identical) to those of humans and whether or not some future organism or some other deterministic machine might eventually reach higher levels of self-awareness are questions for the future. Perhaps humans have reached an adaptive peak in the evolution of consciousness, so further advances might require brains evolved in a new anatomical/physiological direction.

Conclusion

The mathematical analysis of optimal behavior in noisy situations thus indicates that (1) noise is an inescapable component of communication, (2) subjective awareness of self is a higher-order association of perceptions and responses, (3) decision-making is fundamental component of all communication and perception, and (4) both processes are as unpredictable as the unavoidable noise. An advantage of this analysis of the evolution of communication in noise is the framework it provides for addressing the questions posed at the start of this essay.

To account for the source and the content of self-awareness, previous discussion has always relied either on supernatural intervention or on vague neural operations on purified sensations. Supernatural intervention of course obviates any mechanistic explanation, including evolution. Response to pure sensations, on the other hand, leaves each organism encased in its own perceptions, without a way to distinguish between subjective and objective events. The evolution of noisy communication, in contrast, shows that self-awareness (consciousness) results directly from the operations of nervous systems exposed to noisy sensations. We can expect that the neural correlates of self-awareness will depend not only on sensations of interest to an organism but also on the noise mixed with them. The resulting explanation for self-awareness requires no unnatural or unspecified components.

The principal conclusion from such an analysis is that the problems of consciousness might reduce to problems of evolution, signal detection, and neurobiology, all highly mathematical and physical, and thus mature scientific fields. Discussion of the mechanisms of consciousness might therefore migrate from philosophical to scientific discourse.

Cross-References

► Evolution of Communication
► Evolution of Free Will

References

De Veer, M. W., & van den Bos, R. (1999). A critical review of methodology and interpretation of mirror self-recognition research in nonhuman primates. *Animal Behaviour, 58*, 459–468.

Dennett, D. C. (1991). *Consciousness explained*. Boston: Little, Brown and Company.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activities (readiness potential); the unconscious initiation of a freely voluntary act. *Brain, 106*, 623–642.

Lurz, R., Krachun, C., Mahovetz, L., Wilson, M. J. G., & Hopkins, W. (2018). Chimpanzees gesture to humans in mirrors: Using reflection to dissociate seeing from line of gaze. *Animal Behaviour, 135*, 239–249.

Lutz, R. W. (Ed.). (2009). *The philosophy of animal minds*. Cambridge: Cambridge University Press.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press. Reprint, Mahwah: Erlbaum, 2004.

Savage-Rumbaugh, S., Shanker, S. G., & Taylor, T. J. (1998). *Apes, language, and the human mind*. Oxford: Oxford University Press.

Searle, J. R. (1997). *The mystery of consciousness*. New York: New York Review of Books.

Shannon, C. E., & Weaver, W. (1963). *The mathematical theory of communication*. Urbana: University of Illinois Press.

Wiley, R. H. (2015). *Noise matters: The evolution of communication*. Cambridge: Harvard University Press.