Being Rational

PHIL ROBERTS, JR.
philrob@ix.netcom.com

Abstract: Based on a simple premise and relying on the metaphor of vision, I offer an alternative to the "standard picture" of rationality in which 'being rational' is construed as simply a matter of conforming to established rules of inference. In addition to offering a means of visualizing a number of features commonly associated with rationality including correspondence, coherence, logic, incompleteness, justification and irrationality, I also explain how my alternative to the "standard picture" can provide solutions to the Gettier problem, the "rationality debate" (Stein) and "the central theoretical problem of sociobiology" (Wilson).

> *The concept of rationality, one might say, is incorrigibly elusive... I believe it is fair to say that in philosophical discussions of rationality, there is a sense in which we do not "know what we are talking about" and can never do so, if what is demanded is a concise definition* (Max Black).

Assuming that Black is not too far off the mark on this, it would be foolish to assume that one can develop a theory of rationality where so many have failed without having some way of stacking the deck. Or, to carry the metaphor a step further, we might think of this in terms of discovering a 'tell' that Mother Nature has let slip that can give us some indication of the hand she might be holding -- some innocuous feature that others may have overlooked perhaps, that might serve to put the study of rationality on an empirical footing. And, in this regard, I have long been enamored of the premise that 'feelings of worthlessness' are not so much an adaptation as a maladaptive byproduct of the evolution of rationality[1] -- part of the price we humans have had to pay for having become a little too rational/ objective for our own good or -- in terms of the theory of natural selection -- part of the cost of doing business that Mother Nature "tolerates" as a necessary premium for having a rational species to do her bidding.[2]

The gist of the argument here is that, as humans became smarter and smarter about how the world is put together -- presumably transpiring over millennia of linguistic and cultural evolution -- they gradually became smarter and smarter about values. That is to say, contrary to Hume's famous dictum that reason is strictly a matter of truth and falsehood, they gradually became increasingly adept at distinguishing between values that are justifiable (e.g., supported by evidence)[3] and those that are simply a matter of happenstance. Eventually, so this story goes, this culminated in an increased volatility in the most crucial value of all –- one that in all likelihood lies at the very heart of the will to survive –- an increased volatility in self-value along with a host of maladaptive effects (feelings of worthlessness, anxiety, depression, addiction, suicide, guilt, etc.). Expressed in an outline format, the theory I have in mind here would read as follows:

> **Objective:** To answer the question: 'Why are the members of one particular species of naturally selected organism expending significant amounts of effort and energy on the biologically bizarre non-physical objective of maximizing self-worth?'[4]
>
> **Explanation:** Being the blind arational process that she is, Mother Nature instills in all her creatures a sense of their own importance (or of the importance of their needs) that is rationally inordinate. And, as a species reaches a certain stage in its rational/ cultural/ memetic development, its members increasingly come to question this inordinacy, and increasingly come to require reasons (justification) for maintaining it (needs for love, purpose, meaning, moral integrity, autonomy, justice, dignity, etc.).

Incompleteness

The implications of the foregoing conjecture are extensive, including implications with respect to emotional disorder, indeterminism, free will, rationality, ethics and incompleteness:

> **Incompleteness:** When viewed from the context of our underlying premise, 'feelings of worthlessness' can be construed as providing an empirical vindication of the Lucas/ Penrose assertion that Godel's incompleteness theorem (1931) can be construed as an argument that "minds are different from machines":
>
> > *Godel's theorem states that in any consistent system which is strong enough to produce simple arithmetic there are formulae which cannot be proved-in-the-system, but which we can "see" to be true... Godel's theorem must apply to cybernetical machines, because it is of the essence of being a machine, that it should be a concrete instantiation of a formal system. It follows that given any machine which is consistent and capable of doing simple arithmetic, there is a formula which it is incapable of producing as being true -- but which we can "see" to be true. It follows that no machine can be a complete or adequate model of the mind, that minds are essentially different from machines…*(Lucas, 1961).
>
> The empirical vindication of this much contested assertion is based on the following line of inference:
>
> 1. 'Feelings of worthlessness' constitute evidence that humans are beginning to question the value of their existence and therefore are beginning to question the value of achieving the fixed objective of Mother Nature's most basic program (survival).[5]
>
> 2. The same capacity for "standing outside the system" (Lucas) that enables us to "see" that the Godel sentence is "true" is what enables us to stand outside of nature's most basic program and question (in the guise of 'feelings of worthlessness') whether it is one worth completing.
>
> 3. Rationality cannot be constrained (captured in its entirety) within a formal system, in this case, one that has been "designed" to achieve a fixed objective, not even by Mother Nature herself.

Although this argument assumes at least some prior familiarity with the relevant literature (Penrose, Hofstadter, Lucas, etc.), the pertinent point here is that, in a world where my premise is correct and I have not misread its implications, one would have reason to suspect that all concrete instantiations of rationality are likely to be incomplete (e.g., open-ended) and therefore that no individual, culture, belief, theory, value, objective, etc. is likely to be rational in any but a quantitatively relative sense of the term (e.g., X is relatively more or less rational than Y, the norm, etc.). This, in turn, would mean that when we refer to an individual as "rational" or "irrational" that we are probably just expressing a rough appraisal of how this individual's rationality compares to the norm.[6]

Being Rational

If one assumes, as I have conjectured, that feelings of worthlessness are a maladaptive byproduct of the evolution of rationality in the sense of resulting from a more objective understanding of how the world is put together, it also seems to follow that the "standard picture" of rationality is mistaken.[7] 'Being rational' is not so much a matter of slavishly conforming to established rules of inference (a process) as a matter of 'being able to "see" what is going on' as a result of reasoning that has already transpired, whether one's own, or culturally acquired – a state of mind facilitated by the product of a process. This would also mean that, rather than an assessment of one's compliance with established rules of inference, our common sense rationality ascriptions might better be construed as appraisals of a mental map of sorts, one apparently comprised of both beliefs and values, in which the cognitive component of this "seeing" correlates with the extent to which the map is comprised of beliefs that accurately and coherently represent reality including, among many other things, beliefs about how to acquire beliefs that accurately and coherently represent reality, reflected in how well one reasons.

The "Picture-Picture"

Cognitive Component: Since, in this view, 'being rational' constitutes a state of mind facilitated by the product of a process, the most likely candidate for the referent for the term, 'rationality', would be the so-called "map" that facilitates this state of mind. And, in deference to Black's admonition about the futility of attempting to come up with a literal definition,[8] I have found it helpful to simply visualize this psychical product in terms of a follow-the-dots diagram in which "seeing" is facilitated by the extent to which the lines have been connected correctly (correspondence) and in which cohering lines (coherence) count for more than dispersed lines in terms of facilitating the ability to "see" what the diagram represents.

For example, envision a diagram A that represents the face of a well known individual in which there is a significant number of correctly connected lines that are so widely dispersed as to render the diagram unintelligible. Compare this to a diagram B in which there are far fewer correctly connected lines but in which the lines cohere sufficiently for one to be able to decipher a facial feature, such as an eye or a nose, and at least develop a vague idea of what the diagram as a whole represents. In terms of mental "seeing", one would opt for Diagram B in spite of the fact that there are far fewer correctly connected lines (fewer individuated "true" beliefs, facts, observations, etc.) than are present in Diagram A.

Coherence:

> *All kinds of reasoning consist in nothing but a comparison and a discovery of those relations either constant or inconstant, which two or more objects bear to each other* (David Hume).

If Hume is right about this, and if we assume that reasoning is the process that produces the product that facilitates "seeing", then there is reason to suspect that the "glue" that holds the cohering elements of the mental map together (the cohering lines in the diagram) is likely to be analogical (nonlogical) in nature, e.g., a function of the extent to which water flowing in a pipe can indeed serve as an appropriate vehicle for coming to "see" how electricity behaves, e.g., in which there are correlates of pressure (voltage), size of pipe (amperage) and amount (wattage):

> *If analogy were merely a special variety of something that in itself lies way out on the peripheries, then it would be but an itty bitty blip in the broad blue sky of cognition. To me, however, analogy is anything but a bitty blip -- rather, it's the very blue that fills the whole sky of cognition – analogy is everything…* (Douglas Hofstadter).

Logic: If we assume that 'being logical' is simply a matter of conforming to established rules of inference, and in which the term is taken to include everything from probability theory to the logic of discovery, I would argue that, at least where epistemic matters are concerned, the term 'logic' might reasonably be construed as referring to:

> 'the order we have so far managed to cognize in the manner in which we cognize order.'

Assuming this can be loosely construed as another way of talking about "beliefs about how to acquire beliefs that accurately and coherently represent reality" mentioned earlier as a part of the "map" that facilitates "seeing" ('Being Rational'), we would visualize the rationality of one's logical beliefs in terms of the correctness and coherence of the lines in the portion of the diagram that represents those beliefs. Not only does this allow for a visualization of the distinction between my own theory and the "standard picture" (in which 'being rational' equates with 'being logical'), but it also affords a means of visualizing our common sense intuition that, although in some manner related, and contra to the "standard picture", 'being rational' is by no means synonymous with 'being logical'.

Valuative Component: If one represents value in terms of the darkness of the lines, it is also possible to visualize the irrationalizing influence of powerful emotions. This can be accomplished by visualizing a very dark relatively small region in the diagram contrasted against a background of light barely perceptible lines comprising the remainder of the diagram as a means of representing the manner in which a powerful emotion can render it difficult to appreciate "the big picture".

Incompleteness: In addition to the ability to visualize correspondence, coherence, logic and the irrationalizing influence of emotion, my proposed alternative (in which 'being rational' is a matter of 'being able to "see" what is going on relative to the norm') to the "standard picture" (in which 'being rational' is simply a matter of conforming to established rules of inference) can also accommodate the incompleteness feature. This can be visualized in terms of a diagram with no definite boundaries and/ or no limitation on the amount of possible detail.

Justification: What might seem like compelling evidence to you or me might appear to be highly suspect to Sherlock Holmes, which raises an interesting point about the nature of justification (e.g., the rationality of believing that X). As with our use of the bivalent terms, 'rational' and 'irrational', to refer to what in all likelihood are relational evaluations (e.g., X is relatively more or less rational than Y, the norm, etc.), perhaps the bivalent terms, 'justified' and 'unjustified', should also be construed as expressing relational evaluations necessitated by the same underlying peculiarity, i.e., that all concrete instantiations of rationality are likely to be incomplete. If so, then various proposed counterexamples to the justified true belief account of knowledge (e.g., Gettier, 1963) would have to be construed, not as examples of justified true belief that fail to qualify as knowledge, but rather as examples of true belief that compare favorably to the norm on the scale of justification and as such, in which there are no guarantees where matters of justification are concerned.

Representing this peculiarity about our use of the terms, 'justified' and 'unjustified', can be accomplished simply by representing those beliefs that are associated with matters of justification, including the justification of one's logical beliefs, in terms of a subsidiary diagram setting off to the side of the main diagram and in which, as with the main diagram, there are no definite boundaries and/ or no limitations on the amount of possible detail (incompleteness). And the likelihood that the beliefs and values represented by the main diagram are relatively well justified can be visualized in terms of the extent to which the lines in the subsidiary diagram have been connected correctly and coherently.

The Rationality Debate

If our common sense rationality ascriptions are indeed implicitly relational, then experimental evidence that humans routinely violate established rules of inference (e.g., Tversky and Kahneman, 1975) should not be construed as evidence that humans are irrational as necessitated by the "standard picture" and as scores of experts on the subject have appeared to maintain (Open Peer Commentary in Cohen, 1981). Nor would it be necessary to navigate the intricacies of a wide verses a narrow reflective equilibrium or a competence/ performance distinction (Cohen target article) or a lengthy examination of the dissimilarities between linguistic and reasoning competence (Stein, 1996) in order to establish why this might be so. This is because, while all ascriptions are implicitly relational of necessity (incompleteness), they are implicitly relational relative to the norm simply as a matter of custom and convenience. As such, there is nothing to preclude us from inferring from the experimental evidence that expert opinion might be relatively more rational than the norm (in terms of 'being able to "see" what is going on') where certain matters of belief acquisition are concerned (e.g., the Wason selection task, the conjunction fallacy, base rate fallacies, etc.) while at the same time preserving and explaining our apparently unshakeable conviction that, where 'being rational' is concerned, ordinary humans are pretty much the standard. And, in accordance with previous discussion of this matter, this expert opinion could be visualized in terms of a diagram in which there are relatively more correct and cohering lines in the region of the diagram which represents beliefs about how to acquire beliefs that accurately and coherently represent reality.

> *…Webster's informs us that irrational means "not endowed with reason or understanding" and "lacking usual or normal mental clarity or coherence." Although it is not explicitly stated in these definitions, it seems fair to say that "usual" and "normal" are relativized to people. Any notion that humans on the whole are irrational in significant, systematic ways would seem, well, irrational to most of us* (Renee Elio).

Being Irrational

The assumption that expert opinion is relatively more rational than the norm where certain matters of belief acquisition are concerned can be something of a two edged sword. This is because, given our common sense understanding that 'being rational' is by no means synonymous with 'being logical', and our common sense understanding that, come hell or high water, ordinary humans are the standard for 'being rational', it could just as readily be argued that expert opinion has been relatively less rational than the norm with regard to its longstanding love affair with the "standard picture". Indeed, some might even go so far as to suggest that, as is so often the case, perhaps the lover in question has been blinded by lust, on this occasion, the lust to reduce mind to matter via the reduction of rationality to logic, rules, processes, principles, procedures, etc. that can be instantiated in a computer.

Assuming, for the sake of demonstration, that this has, indeed, been the case, we would want a way to visualize this feature (wishful thinking) in which one is being less rational as a result of the presence of something as opposed to the absence of something, as in the previous example where the extreme would be arational or non-rational rather than irrational. This could be accomplished simply by visualizing a small dark region against a barely perceptible remainder of relatively well connected lines in the subsidiary diagram where 'being able to "see"' the explanatory shortcomings of the "standard picture" has been compromised, not by a deficiency in understanding where matters of justification are concerned, but rather by a lack of valuative objectivity/ impartiality in coming to appreciate those explanatory shortcomings and assess them from a level playing field, so to speak. And the deleterious effect this has had on the beliefs represented by the main diagram could be visualized in terms of incorrectly connected cohering lines representing the false belief (presumably) that the computer is an appropriate vehicle for coming to understand the mind. We might even embellish this visualization a bit in terms of a very dark as opposed to a very light dot connecting these cohering lines to represent the presence of too much faith in the coherence of this particular approach to psychology (computationalism, cognitive "science", homuncular functionalism, etc.) and its manifestation, the "standard picture".

> *It is wrong always, everywhere, and for everyone to believe anything upon insufficient evidence* (W. K. Clifford).

Summarizing: In addition to the ability to visualize correspondence, coherence, logic, incompleteness, justification and irrationality, I have also shown how my proposed alternative to the "standard picture" can provide solutions to the Gettier problem and the rationality debate, both of which have figured prominently in the philosophical literature. In what follows, I will conclude with one final explanatory advantage of the theory, this time within the scientific domain -- its ability to provide an answer to "the central theoretical problem of sociobiology":

> *For decades biologists complacently had thought that selection shapes traits that benefit groups and species (Wynne-Edwards 1962). This assumption made it easy to view self-sacrifice for the sake of the group as entirely natural and expected. With simple but ruthless logic, Williams showed that selection at the group level is feeble compared to selection at the individual level (Williams 1966). Natural selection, it turns out, acts mainly to benefit genes and individuals, not groups or species (Maynard Smith 1964). Many implications follow from this, but the most profound is the transformation of altruism from a natural tendency into an evolutionary mystery (Dawkins 1976; Badcock 1986; Barash 1977; Krebs 1970). E. O. Wilson called altruism "the central theoretical problem of sociobiology" (Wilson 1975, 3).* (Randolph Nesse)

Being Good

Perhaps the most intriguing implication of my basic premise (that feelings of worthlessness are a maladaptive byproduct of the evolution of rationality) lies in the domain of ethics:

Ethics: Since, according to this explanation, 'being more rational' correlates with 'being more valuatively objective/ impartial', the moral maxim, 'Love (intrinsically value) your neighbor as you love (intrinsically value) yourself' could be construed as an imperative of an implicit theory of rationality in which 'being rational' entails (among other things) 'being valuatively objective/ impartial'. This would also mean that, to the extent this "implicit theory" turns out to be "true", the author of Genesis actually got it right in referring to our awareness of right and wrong as a form of knowledge (moral realism).

Moral Sentiment: If the emergence of an awareness of right and wrong is simply an emerging awareness of the nature of rationality itself, then it would also provide a rationale for the persuasive force of moral argument, in that it could be construed as exploiting our need to view ourselves as rational as a determinant in assessing self-worth. Although, as a part of the basis for assessing self-worth, none of us can actually measure up to the standard of loving others as we love ourselves (valuative objectivity/ impartiality), in this view we nonetheless come to experience feelings of worthlessness (guilt) along with a corresponding reduction in the will to survive (depression) when we deviate from the standard to an unacceptable degree. In other words, a capacity for guilt (having a conscience) is a part of the price we humans have had to pay for having become a little too rational/ objective for our own good – a maladaptive (from the standpoint of the individual or "gene") manifestation of our need to justify our existence by conforming (to a reasonable degree relative to the norm) to a shared subconscious "theory" of rationality in which 'being rational' entails (among other things) 'being valuatively objective/ impartial'.

We struggle to be honest and courteous and responsible and brave in circumstances where it is difficult. Even if apes are sometimes courteous, responsible, and brave, it is not because they think they should be. Even as primitive a phenomenon as a teenager's efforts to be "cool" is a manifestation of the human tendency to live a life guided by ideals rather than merely driven by impulses and desires. We also suffer deeply from our self-evaluations and act in sick and evil ways as a result. This is part of what I had in mind earlier when I said that human beings seem psychologically damaged in a way that suggests a break with nature (Christine Korsgaard).

We are nicer than what is good for our selfish genes (Richard Dawkins).

Endnotes

1. In contrast with those who equate naturalism with materialism, I have never thought there was anything particularly "scientific" about the assumption that the mind/ body problem has somehow been miraculously resolved as a result of admittedly amazing advances in the material sciences. As such, it should come as no surprise that I am assuming that 'feelings of worthlessness' are a non-physical feature of nature – one produced by millennia of linguistic and cultural evolution operating in tandem with millions of years of neurological evolution – that are probably present only in enculturated human minds. The wild boy of Aveyron and, prior to her famous Aha!, Helen Keller, no doubt experienced lots of feelings (pain, fear, aggression, frustration, despair, etc.), but I suspect that feelings of worthlessness were probably not among them.

2. In somewhat more literal terms, the maladaptiveness of an increased volatility in the will to survive (feelings of worthlessness, suicide, etc.) is outweighed by the adaptiveness of a massive increase in the ability to survive (agriculture, technology, global relief, etc), both of which are, in this view, facilitated by an increase in understanding.

3. Hume, himself, offers an example of "objective evidence" concerning one's worth in Book III, Part I, Section XI of his *Treatise of Human Nature*:

   > *Tis now time to turn our view from the general consideration of sympathy, to its influence on pride and humility, when these passions arise from praise and blame, from reputation and infamy..... Now nothing is more natural than for us to embrace the opinions of others in this particular; both from sympathy, which renders all their sentiments intimately present to us; and from reasoning, which makes us regard their judgment, as a kind of argument for what they affirm* [my underline]. *These two principles of authority and sympathy influence almost all our opinions; but must have a peculiar influence, when we judge of our own worth and character....*

4. This should not be conflated with the assertion that humans are expending significant amounts of effort and energy on the conscious pursuit of self-worth, but merely the claim that the need is ubiquitous even if not always consciously appreciated:

   > *There is no value-judgment more important to man -- no factor more decisive in his psychological development and motivation -- than the estimate he passes on himself. This estimate is ordinarily experienced by him, not in the form of a conscious, verbalized judgment, but in the form of a feeling, a feeling that can be hard to isolate and identify because he experiences it constantly: it is part of every other feeling, it is involved in his every emotional response. ... it is the single most significant key to his behavior* (Nathaniel Branden).

5. Technically, it's a matter of surviving long enough to successfully reproduce and with "special" and "limited" exceptions (Dawkins, 1976) in deference to inclusive fitness theory (Hamilton, 1964).

6. I do not take departures from the ordinary use of words lightly. However, I also do not construe ordinary usage as sacrosanct, particularly if a plausible rationale can be offered for why it might be mistaken such as that relational terminology is simply too cumbersome to be employed for the purposes of every day conversation.

   > *Our common stock of words embodies all the distinctions men have found worth drawing, and the connexions they have found worth marking, in the lifetimes of many generations: these surely are likely to be more numerous, more sound, since they have stood up to the long test of the survival of the fittest, and more subtle, at least in all ordinary and reasonably practical matters, than any that you or I are likely to think up in our arm-chairs of an afternoon* (J. L. Austin).

7. Quoting Stein:

   > *I call the claim that humans are rational the* rationality thesis *and the claim that humans are irrational the* irrationality thesis. *Both of these claims are typically based on what I call* the standard picture of rationality.

According to this picture, to be rational is to reason in accordance with principles of reasoning that are based on rules of logic, probability theory, and so forth (Stein, 1996, p. 4).

8. If called upon for a literal definition for the term "rationality" given the implications of my underlying premise, the devil is most definitely in the details. As a first approximation, I would opt for 'the psychical (non-physical) product of ampliative inference', most of which is acquired from one's culture, and with the term "ampliative" unpacked to refer to inferences that serve to maximize the objectivity, understanding, impartiality, wisdom, etc. of the agent or agents so endowed. However, since we usually employ the term in an evaluative context, it gets messy, particularly when you take into consideration that these evaluations or assessments are relational in nature (incompleteness). In this sense, are you ready for this, the term refers to 'the extent to which one's psychical product of ampliative inference serves to maximize one's objectivity (one's mental ocularity) relative to the norm'.

References

1. Cohen, L. Jonathan (1981), 'Can Human Irrationality Be Experimentally Demonstrated?', *Behavioral and Brain Sciences*, 4: 317-370; 6: 487-533; 10: 311-13.
2. Dawkins, Richard (1976), *The Selfish Gene*.
3. Gettier, Edmund (1963), 'Is Justified True Belief Knowledge?', *Analysis* 23, pp 121-3.
4. Godel, Kurt (1962), *On Formally Undecidable Propositions* (translation of 1931 paper).
5. Hamilton, W. D. (1964), 'The Genetical Evolution of Social Behavior', *Journal of Theoretical Biology*, 7: 1-16; 17-52.
6. Hume, David (1739), *A Treatise of Human Nature*.
7. Lucas, J.R.. (1961), 'Minds, Machines and Godel', *Philosophy*, 36, p. 112.
8. Stein, Edward (1996), *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science*.
9. Tversky, Amos and Daniel Kahneman (1974), 'Judgment Under Uncertainty: Heuristics and Biases', *Science* 185: 1124-31.