

Truth as an Attractor in Peirce's Evolutionary Epistemology

Abstract

Have we evolved an inherent tendency to believe truths over falsehoods? Given our variety of powerful cognitive biases, it seems unlikely that we have any predisposition toward truth. However, truth might still be a *weak attractor*, which doxastic agents and communities would evolve toward over the long run. This view is upheld by Charles S. Peirce, on whose “pragmatic maxim” a true belief is one that *would* become permanently “fixed” with sufficient time, due not simply to the use of correct methods but also due to *the nature of belief* itself. If his psychological model is plausible, Peirce offers an evolutionary epistemology that cuts across Bradie's (1986) EEM/EET distinction and across descriptive and normative projects, providing a description of how we know things as well as a prescription for knowledge: select methods that would best expedite our approach toward a state of permanent doxastic fixation: the “final opinion”.

1. *Introduction*

As it developed in the 1970s-80s, evolutionary epistemology (EE) came to encompass both normative epistemologies promoting theory variation and a sort of “natural selection” among theories (e.g. Popperian falsificationism) and broadly naturalized epistemologies that seek to understand how various cognitive structures/mechanisms evolved. Bradie (1986) thus distinguishes between two types of EE projects: “evolutionary epistemology of theories” (EET) and “evolutionary epistemology of mechanisms” (EEM). EET theories attempt to model processes of scientific advancement after processes of phenotypical development in biology—with both involving variation, selection, and transmission (of theories, or of phenotypes)—while EEM theories attempt to understand how evolutionary processes led to the biological structures that produce knowledge (primarily in humans, but also in other animals).

Both EEM and EET theories have been subject to much criticism. Thagard (1980) criticizes EET theories as superficial, as the similarities between scientific and biological processes lie not much deeper than there being some type of variation, selection, and

transmission in each respective process. EEM theories often fail to determine *whether* and *for what* a given cognitive structure/mechanism is an evolutionary adaptation, and whether that cognitive structure or mechanism reliably tracks *truth* or produces true beliefs (Sage, 2004). Campbell's BVSR model of knowledge (1960) cuts across Bradie's EET/EEM distinction by postulating evolved cognitive mechanisms that themselves perform variation and selective-retention functions in individual knowers. Campbell's model faces many of the same criticisms, another being that it fails to adequately describe the nature of the "fit" of beliefs or theories with the environment, or how that fit can be understood in terms of *truth*.

It's well known that the early pragmatists—Charles S. Peirce, William James, and John Dewey—drew heavily from and emphasized ideas of biological evolution in their writings. The EE theory that we find in Peirce, which I will sketch here, puts the notion of truth front and center. Like Campbell's program, it cuts across Bradie's EET/EEM distinction by treating the selective processes by which knowledge would advance as extensions of evolved cognitive dispositions. But one advantage of Peircean EE is that it centrally involves a theory of truth as the ideal "fit" between our beliefs and reality. Further, his account describes truth as an *attractor state of the mind*. In dynamical systems, an "attractor state" is one towards which a system tends to evolve. On Peirce's view, despite all our biases and limitations, we are cognitively predisposed toward holding true beliefs over false ones, even if only very weakly, such that *in the long run* we would attain and maintain the truth—a state which Peirce calls "the final opinion" (W3:274). The normative project, then, concerns methods of *expediting* this natural process by which false beliefs tend to be selected out while true beliefs tend to be retained. What Peirce calls "the method of science" is broadly the method of accelerating the selection process.

In sketching this model, I will address the Peirce scholarship where necessary. To my knowledge, Skagestad (1979) was first to explicitly attribute an EET theory to Peirce, correctly locating central parts in Peirce's seminal 1877 essay "The Fixation of Belief" (FOB). However, Skagestad doesn't seem to recognize how, for Peirce, truth is a natural attractor. It seems generally presumed by scholars that truth as the final opinion can be approached only if *the correct methods* are employed, where the employment of such methods is not inevitable. To the contrary, on Peirce's account, the adoption of correct methods may be either unnecessary or inevitable with sufficient time.

2. *Belief, Perception, Doubt*

A major premise of Peirce's EE, stated explicitly in FOB, is that *doubt* is an "irritating" state that compels one "to attain a state of belief" (W3:247). This idea isn't original to Peirce, who likely drew it from Alexander Bain,¹ and it seems broadly supported by more recent work in cognitive dissonance theory showing that people commonly experience cognitive dissonance as distressful and that they seek to reduce it typically by means of some attitudinal or doxastic change.² It's not uncommon for people to have dissonant beliefs and fail to experience any "irritation of doubt", especially if the beliefs are context-sensitive or their inconsistencies lie in their logical consequences (since belief is not closed under straight entailment). However, when made fully aware that our beliefs are contradictory, our common response is some attempt at attitudinal or doxastic change. Peng & Nisbett (1999) showed that the type of doxastic change attempted is culturally sensitive, with East Asians preferring to find "middle ways" between contradictory belief and Westerners preferring to surrender one belief for the other.

¹ See Murphey 1961, p.161

² Harmon-Jones and Mills 1999

Peirce isn't too clear that doubt involves dissonant beliefs until around 1903, when he develops his account of *perception*. On this account, the perceptual process is involuntary until it terminates in a *perceptual judgment*, or the mental assertion of a proposition in response to some perceptual experience or "percept".³ Although the perceptual judgment marks the point at which we can begin controlling our thought concerning what we perceive, the perceptual judgment itself is uncontrollable.⁴ And often equally uncontrollable is *the recognition that a perceptual judgement contradicts a proposition we expected to be true*. Such recognition then causes doubt or dissonance.

For Peirce, a primary function of belief is set to expectations, such that even beliefs about past events set expectations about conditional futures. This is a consequence of the theory which Peirce is best known for: *pragmatism* (or, after 1905, *pragmaticism*). Though sometimes he claims that it's only a method of analysis, *pragmaticism* is more substantially the theory that, beyond the formal significance of their verbal definitions, the meaning of any *general term* or any predication thereof is best represented by a series of action-perception/perception-action conditionals in the subjunctive mood.⁵ For instance, to predicate "hard" to a subject is to signify that if one were to press upon the subject, it would resist deformation. Thus, upon *pragmaticism*, any predication of a general term to some subject consists in possible conditional expectations of perception. If one *believes* the proposition, then one is *disposed to have* those conditional expectations.

Crucially, for Peirce, these subjunctive action-perception conditionals are not analytic to a general term. They're representations of interconnections among concepts. His *pragmatism*

³ For Peirce on "the percept", see CP 2.142-46, 1902 and CP 7.619-644, 1903.

⁴ CP 5.55 and 5.116-142, 1903.

⁵ CP 5.457, CP5.480-6, and EP2:241.

involves a form of *meaning anti-atomism*.⁶ While the meaning of some general terms might involve nothing further than the formal features of its verbal definition (what Peirce calls “the second grade of clearness”), Peirce is confident that the meaning of most general terms lies in their significance for perception and conduct. Thus, almost any belief involves perceptual expectations, and so almost any belief can involve expectations that can be “disappointed” or contradicted by a perceptual judgment. However, given the anti-atomistic nature of pragmatic meaning, there may be numerous beliefs whose pragmatic meaning includes a given expectation. When the expectation is disappointed, the doubt or dissonance may be distributed evenly or unevenly, shaking up true beliefs as much as false ones. Though in particular cases this may result in the removal of true beliefs instead of false ones, on Peirce’s view, beliefs whose pragmatic meaning consists mostly in expectations that would eventually be contradicted by perceptual judgments *tend to be the false ones*. To understand why, we must turn to Peirce’s account of truth.

3. *Truth, or Permanently Settled Belief*

While both true and false beliefs are subject to being “selected out” in any given mind by sufficiently strong doubts over some definite period of time, *in the long run*, false beliefs are subject to permanent *removal* while true beliefs are subject to permanent *settlement* or *fixation*. Peirce does not merely stipulate these claims about *truth* and *falsity*. They are consequences of Peirce’s pragmatic clarifications of the concept of *truth*, based on his theory of pragmatic meaning.

In “How to Make Our Ideas Clear” (1878), after establishing the verbal definition of *reality* as “that whose characters are independent of what anybody may think them to be”

⁶ EP2:305, 1904.

(W3:271), he argues that the main significance of reality for perception is its propensity to determine belief *via* perception. The “experiential method of settling opinion”, which Peirce sketches in FOB, seeks to maximize the potential of reality to determine our beliefs. Anyone who follows this method, then, would *in the long run* be equally determined to believe certain propositions over others. So, our clearest conception of *reality*, Peirce argues, is of something that would sooner or later determine belief or consensus *about it*. Hence, “the opinion which is fated to be ultimately agreed to by all who investigate, is what we mean by the truth, and the object represented in this opinion is the real” (W3:273). Or, as he puts it years later, truth is “the predestined result to which sufficient inquiry *would* ultimately lead” (CP5.494).

On some interpretations, truth as “the final opinion” is primarily a regulative hope of inquiry,⁷ where we can hope to achieve a final consensus on some matter but it might be unlikely that we will (after all, civilization could end before that happens). But truth as the final opinion is not a matter of what actually will happen but *what would happen with sufficient time*; and although Peirce does express the final opinion as a hope, it seems clear he thinks there is a real, if even fairly weak, propensity for truths to get believed, while not so for falsities:

For truths, on average, have a greater tendency to get believed than falsities have. Were it otherwise, considering that there are myriads of false hypothesis to account for any given phenomenon, against one sole true one (or if you will have it so, against every true one), the first step toward genuine knowledge must have been next to impossible. (CP5.431, 1905)

Passages such as this one show that, for Peirce, truth as the final opinion is more than a hope: *it is statistically eventual outcome*. Further, while his pragmatic clarifications of *truth* might be read as saying only that true propositions would be believed *at some point or other*, and not

⁷ E.g., Misak 1999 and Hookway 2000

that those beliefs would also become *permanently settled*, that they would also become permanently settled is shown in FOB, where “the experiential method” is presented as the only method that could result in *permanently settled beliefs*, remaining free of doubt. Peirce is clear that “to believe the absolute truth would be to have such a belief that under no circumstances, such as actually occur, should we find ourselves surprised” (R 693:166, 1902).

Peirce refers to “absolute truth” in passage quoted above perhaps to emphasize that *no part* of the true belief—as in *not one* of the conditional expectations in which its pragmatic meaning consists—is subject to disappointment. However, as indicated earlier, any such conditional is implicated in many different beliefs, each of which would also have to be “absolutely true” in order for any of them to remain totally free of doubt, given other inferential relationships between the beliefs. This means that beliefs could not approach permanent settlement in isolation of each other, but only as a whole doxastic network. This should be expected given Peirce’s meaning anti-atomism, mentioned earlier: truth isn’t a property that a belief can have separately from other beliefs.

Many find Peirce’s pragmatist analysis of truth objectionable because it seems likely there are true propositions on which belief would never be settled, as all evidence of such truths may be inaccessible to humans. In Peirce’s scholarship, this is known as *the problem of buried secrets*. Peirce’s main response is optimism. He remarks: “Give science only a hundred more centuries of increase in geometrical progression and she may be expected to find that the sound waves of Aristotle's voice have somehow recorded themselves.” (CP 5.542, 1902) Some might object that Relativity demonstrates that there are unknowable truths, as there seem to be facts beyond our own past and future *light cones* that we could never know. But Relativity only limits causality and not knowledge. We might discover how to reliably *infer* distributions of

matter and energy beyond our light cones from the information contained within them. (Moreover, given the range of physical phenomena that remain unexplained, Relativity could still turn out false.) However, Peirce need not be read as claiming that the *only* true propositions are the ones that would become permanently fixed beliefs. Peirce accepts the *verbal definition* of truth as *correspondence with reality* (e.g., W3:282, 1878). The pragmatist analysis of truth is an explanation of *what this correspondence could mean* in terms of perception and conduct. However, there might be propositions whose truth cannot be pragmatically clarified and only be verbally defined. In that case, Peirce's pragmatic theory of truth need not be treated covering *all* legitimate senses of "truth".⁸

4. *Belief Generation and "Logical Fatalism"*

So far I have explained the *selection* and *transmission* features of Peirce's EET, which involve three key theses: (a) any belief has semantic relationships to perceptual expectations; (b) doubt or dissonance caused by unsuccessful expectation are primary causes of belief removal; and (c) via Peirce's pragmatism, *truth* is understood in terms of successful expectation and avoidance of doubt while *falsity* is understood in terms of unsuccessful expectation and proneness to doubt. However, it hardly makes sense to analyze *truth* in terms of permanently fixed belief unless we assume a mechanism for belief *generation*, such that any true proposition would eventually *occur to one to believe*.

What Peirce calls *abduction* is often seen as a hypothesis-generating inference type, where true hypotheses would eventually become beliefs and where involuntary thought-generating processes, such as the process that generates perceptual judgments, are considered analogous to abduction (CP5.184-85). However, for Peirce, there is a more basic belief-

⁸ See Lane 2018 for an extensive discussion of and proposed solutions to the problem of buried secrets.

generating feature that he calls “the law of mind”. He writes: “the highest quality of mind involves a great readiness to take habits, and a great readiness to lose them” (CP6.613, 1891). Peirce follows association-psychologists in using “habit” to describe the main type of cognitive mechanism—beliefs being *habits of expectation*—and he recognizes, as many of the associationists had, the great plasticity of the mind, allowing it to lose habits but also to spontaneously generate new ones. Peirce conjectures that ideas “naturally ally” themselves with each other—a process he calls “association by resemblance” although he argues that the association creates the resemblance, and not vice versa (CP4.157, 1897)—which then through other associative processes form general ideas and mental propositions (CP7.498-99). Not all mental propositions will lead to beliefs, let alone true beliefs; but in time all true propositions humanly conceivable would actually be conceived and subsequently believed—not because with enough time all physical possibilities be realized, but because of actual predispositions of human cognition.

There is much in Peirce’s cognitive psychology that requires additional empirical support; however, I haven’t found much evidence clearly contradicting his conjectures concerning belief, doubt, perception, and mental plasticity, which, combined with his pragmatist analyses of meaning and truth, suggest that truth is a natural attractor. So it seems plausible enough to motivate further research. Peirce scholars, however, might dispute that Peirce always regarded truth as a natural attractor, though it seems clear that he did early on. In 1871 he claims: “There is, then, to every question a true answer, a final conclusion, to which the opinion of every man is constantly gravitating. He may for a time recede from it, but give him more experience and time for consideration, and he will finally approach it” (CP8.12). Peirce uses the gravitational metaphor at least once more in 1873 (CP7.344).

However, Aktins (2018) has recently argued that while early on Peirce held truth to be the “predestinate opinion”, later (in the 1900s) he held it to be approachable only on the condition that we obtain *the correct methods* of inquiry. This is suggested by such passages as the following; Peirce writes: “if we can find out the right method of thinking and can follow it out, the right method of transforming signs, then truth can be nothing more nor less than the last result to which the following out of this method would ultimately carry us” (EP2:380, 1906). This passage can be understood consistently with the claim that Peirce continued to hold truth to be predestined, if he held that obtaining “the right method of thinking” and “following it out” *are themselves predestined*.

In an unpublished piece written around the same time (1905), Peirce argues for what he there calls “logical fatalism”, which says that “[s]ooner or later [one] will attain the truth,” where “if you take the most pigheaded and passionate of men who has sworn by all the gods that he will never allow himself to believe the earth is round, and give him time enough, and cram that time with experience in the pertinent sphere, and he will surely come to and rest in the truth about the form of the earth” (CP7.78/R 603, c.1905). Perhaps Peirce thinks that, with enough time, this man would eventually come around to the “right method of thinking”. However, he goes on to give an analogy involving a derelict ship floating about the ocean, which would eventually reach some shore on its own, suggesting that the right method is *not* necessary. He further argues that “a vessel which should go and take that derelict in tow ...would be ‘abbreviating’ or ‘expediting’ the fulfillment of the destiny of that derelict in the same sense in which I hold that logic ‘abbreviates’ inquiry, and ‘expedites’ its result. ...This is the sense in which logic ‘abbreviates’ and ‘expedites’ the attainment of truth.” (ibid.)

Correct methods are necessary only to help ensure that truth, or the final opinion, is attained within the time we have as a civilization. As Peirce views the natural process that tends toward truth as already involving variation, selection, and transmission, presumably the correct methods would likely enhance or accelerate these steps. Indeed, one might find that these steps map onto Peirce's distinction between *abductive*, *deductive*, and *inductive* steps of inquiry, which are central to the normative part of his epistemology. But that distinction must be saved for discussion.

Works Cited

- Atkins, Richard K.. "Peirce on Truth as the Predestinate Opinion." *European Journal of Philosophy* 26:1, 411-429.
- Bradie, Michael. 1986. "Assessing Evolutionary Epistemology." *Biology & Philosophy* 1, 401-459.
- Campbell, Donald T. 1960. "Blind Variation and Selective Retention in Creative Thought as in Other Knowledge Processes." *Psychological Review* 67(6): 380-400.
- Harmon-Jones, Eddie & Mills, Judson. 1999. "An Introduction to Cognitive Dissonance Theory and an Overview of Current Perspectives on the Theory." *American Psychological Association*, 3–21.
- Hookway, Christopher. 2000. *Truth, Rationality, and Pragmatism: Themes from Peirce*. Oxford: Oxford University Press.
- Lane, Robert. 2018. *Peirce on Realism and Idealism*. Cambridge: Cambridge University Press
- Misak, Cheryl. 1991. *Truth and the End of Inquiry: A Peircean Account of Truth*. Oxford: Clarendon Press
- Murphey, Murray. G. 1961. *The Development of Peirce's Philosophy*. Cambridge, MA: Harvard University Press.
- Peng, Kaiping & Nisbett, Richard.E. 1999. "Culture, Dialectics, and Reasoning about Contradiction." *American Psychologist*, 54.9, 741-754.
- Peirce, Charles S. *The Collected Papers of Charles Sanders Peirce* 8 vols. (CP) Edited by Charles Hartshorne, Paul Weiss, and Arthur W. Burks. Cambridge, MA: Harvard University Press, 1931–1958. References to the *Collected Papers* are denoted only by the volume number followed by the paragraph number.
- *The Essential Peirce*, 2 vols. (EP). Edited by Nathan Houser, Christian Kloesel, and the Peirce Edition Project. Indiana University Press, Bloomington, Indiana: Indiana University Press, 1992, 1998.

----- *The Charles S. Peirce Papers* (R). Houghton Library, Harvard University. Reference is to the system of categorization in R.S. Robin, *The Annotated Catalogue of the Charles S. Peirce Papers*, and R.S. Robin, *The Peirce Papers: A Supplementary Catalogue*.

----- *Writings of Charles S. Peirce: A Chronological Edition*, 8 vols. (W). Edited by the Peirce Edition Project. Bloomington, IN: Indiana University Press, 1982, 1984, 1986, 1989, 1993, 2000, 2010.

Sage, James. 2004. "Truth-Reliability and the Evolution of Human Cognitive Faculties." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 117(1/2), 95-106.

Skagestad, Peter. 1979. C. S. Peirce on Biological Evolution and Scientific Progress. *Synthese*, 41(1), 85-114.

Thagard, Paul. 1980. Against Evolutionary Epistemology. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, 1980*, 187-196.