Fallibilism is Omega-inconsistent

O Falibilismo é Ômega-inconsistente

T. L. Short
tlshort@earthlink.net

Abstract: Peirce's fallibilism is generally lauded. Little attention has been paid to its appearance of being self-contradictory. That appearance is nonetheless not easily removed. I make an effort to minimize its impact by drawing an analogy of it to ω-inconsistency, in contradistinction to simple inconsistency. But that stratagem proves to be illuminating only by way of revealing the limitations of formal analyses of belief.

Keywords: Belief. Consistency. Fallibilism. Ω-inconsistency.

Resumo: *O falibilismo de Peirce é geralmente louvado. Pouca atenção tem sido dada ao fato de que parece ser autocontraditório. Tal aparência, não obstante, não é facilmente removida. Faço um esforço para minimizar seu impacto, traçando uma analogia entre ela e a ω-inconsistência, em contraposição à inconsistência simples. Mas esse estratagema se prova iluminador somente ao revelar as limitações de quaisquer análises formais da crença.*

Palavras-chave: *Crença. Consistência. Falibilismo. Ω-inconsistência.*

Is fallibilism consistent? Does it matter if it is not?

A.

Consistency, I suggest, is over-valued by philosophers, probably because their work normally lacks fecundity. Physicists and biologists and other scientists work with theories that are imperfect in many ways, often including formalisms that do not always yield calculations that come to the right conclusions. There may even be contradictions that must be placed gently and discretely aside, for the time being. Of course those are defects. But contrary to folks who think that the theory of natural selection is "refuted" because flaws are found in it (a simplistic application of Karl Popper's ideas that Popper himself eschewed), a theory in the natural sciences will be clung to and applied and worked on, despite its grounding false predictions or being unable to explain all that it should explain. It will be clung to as long as it continues to prove fruitful in guiding research and no rival better in that respect emerges. Its very inadequacies are a guide to research: given that in *some* respects it gives us satisfactory results, its inadequacies are then what research seeks to overcome. Inconsistencies can be tolerated if one has some sense, as yet unprincipled but nonetheless real, of where the one part of the theory is

usefully applied and where the other part of the theory, inconsistent with the first, is usefully applied. We hope that the inconsistencies can be corrected somehow. In the meantime, we live in an imperfect world but one that is a lot of fun. In the natural sciences, fecundity trumps consistency, though it shows only nearness to truth, of being on the road to truth.

Not only in science, but also in practical, everyday life we live with inconsistency all of the time. We have an untold number of beliefs, many of which are mutually contradictory. Logicians tell us that this is disastrous. For, from contradictory premises, we may deduce any proposition whatever.[1] Thus, a person with inconsistent beliefs can believe anything. But that disastrous result does *not* occur, for the simple reason that people who are not logicians have the good sense not to draw every conclusion they might draw.

Consider a parallel case. There are men whom other men would trust with their lives but not with their wives. But what if you were to place such a man in a position where he has to choose between your life and your wife? The obvious answer is that you are careful not to place him in such a position. So also, beliefs we know to be contradictory – beliefs that we act on in distinct sorts of circumstances – we are careful not to combine as premises for an inference.[2]

Does that seem impossible? But it is a fact of life! For certain purposes, we believe that p, for other purposes we don't. We realize that we ought to be able to formulate p in such a way that it is limited to the kinds of cases wherein we think it is reliable. But we do not as yet know how to so qualify it. And we are too busy doing other things to make the effort. We try to get along with beliefs that we know are imperfect, just as we try to get along despite our own imperfections.

G.K. Chesterton wrote, "If you argue with a madman, it is extremely probable that you will get the worst of it; for in many ways his mind is all the quicker for not being delayed by the things that go with good judgment. He is not hampered by a sense of humor or by charity or by the dumb certainties of experience..." Good judgment keeps us from drawing all the conclusions we might logically draw.

B.

Now, Peirce's fallibilism appears to entail a contradiction. Each of us has an untold number of beliefs. As a fallibilist I believe that some of my beliefs are false. Of course, I do not know which ones are false. It just seems to me overwhelmingly likely that I am not right about everything. Some of my friends may be surprised by this admission, but that just

---

[1] If from a set of premises and an hypothesis that p is false we are able to derive a contradiction, then we may infer that p is true. That is a *reductio ad absurdum*. It follows that if our premises are themselves contradictory, we may infer any proposition we wish.

[2] In recent years, this has been recognized even by some logicians – those who work on "paraconsistent logics," i.e., systems in which contradictory propositions are allowed but deductions from contradictory premises are blocked. But whether formalization of good judgment is either desirable or possible is another question.

shows that *they* are not right about everything. However, Peirce said, correctly, that "we think [i.e., believe] each one of our beliefs to be true" and that "it is a mere tautology to say so" (EP1:115); indeed, it is redundant. Thus, a fallibilist appears to have contradictory beliefs: he believes that each of his beliefs is true and that some are false. Must one, in order to avoid contradiction, cease being a fallibilist and become a dogmatist?

Perhaps we can evade the problem by weakening the statement of fallibilism. Instead of saying that some of my beliefs are false, I might say that it is possible that some of my beliefs are false, i.e., that I do not know for sure that all are true, that I lack sufficient evidence to be absolutely sure of each and all. In other words, I admit the possibility that continued inquiry may reveal facts that would cause me to alter my beliefs or that would cause some future inquirer to dispute those propositions that I believe. That is a reasonable view to take; indeed, it is a view that it would be unreasonable not to take. And it is consistent with continuing to hold those beliefs – already formed on the basis of experience – that have not yet been disconfirmed by further experience. It would be unreasonable to jettison well-founded beliefs merely because they might some day be disconfirmed. This gambit nevertheless fails, for the following reason.

Peirce identified truth with those propositions that would continue to be affirmed by the community of inquirers at the ideal end of an unlimited inquiry. To think that one's beliefs are true is therefore to think that they will continue to be affirmed at the ideal end of inquiry. Such a thought is consistent with recognition that the propositions believed are logically contingent and, hence, that they are subject to experiential tests. They are propositions of a kind that *can* be disconfirmed, i.e., their being disconfirmed is a logical possibility. But to believe that they *are* true is to believe that they *will not* be disconfirmed *in fact*. Hence, there is still a contradiction between believing my beliefs to be true and believing that they *might* some day be disconfirmed. For that "might" expresses something stronger than logical possibility.

This can perhaps be made clearer by noticing that Peirce's theory of truth is a prediction theory, much like the notorious prediction theory of law due to Justice Oliver Wendell Holmes, Jr. Holmes wrote that an interpretation of what a law means is a prediction about how courts will apply it. That has seemed to many to be deeply cynical, as if it made law to be arbitrarily interpretable, subject to the whims of judges, whose whims lawyers and legal scholars are supposed to try to anticipate. I think, to the contrary, that it is not cynical, that it is, instead, very subtle and is good evidence that there really was a Metaphysical Club in which Holmes and Peirce, in their youth, both participated. Obviously, a judge should not ignore the clear meaning a law has; but that clear meaning goes only so far, and to apply the law correctly to concrete circumstances requires ... judgment! Judgment is based on experience. To say what the law really means is therefore to make a prediction about how it would be applied by honest, intelligent, and earnest judges in light of their experience. Similarly, to say what you think is true is to say what you think future inquirers would continue to affirm – not arbitrarily, but, like you, doing their best to base belief on the evidence.

But if assertions of truth are predictions of what will be affirmed by future inquirers, then they are predictions that disconfirming evidence *will not* be found. And that contradicts any assertion that disconfirming evidence *might be* found. Hence, there is a contradiction between having a belief and thinking that it might prove to be false. If you believe it, you believe that it will not prove to be false.

Try another way of ridding fallibilism of inconsistency. Suppose that one is justified in believing that p when the evidence at hand makes p 99% certain. Then, while I may be justified in believing p and justified in believing q, I might not be justified in believing p&q (the conjunction of p and q, which is true if and only if p and q are both true). For if the evidence makes each of p and q just barely 99% certain, then the evidence makes both together only .99´.99 = .9801 certain, which is not enough to justify believing p&q, if justification requires 99% certainty. And if one has 500 beliefs, each of which is 99% certain, then the conjunction of all 500 is less than 1% certain ($.99^{500}$ = .00657). In fact, it is more than 99% certain that not all 500 are true. So, if you have that many beliefs, you are justified in having one more belief, which is that at least one of your beliefs is false.

That reasoning may reveal something about the logic behind fallibilism, but it does not eliminate fallibilism's inconsistency. For the reasoning pertains to belief's justification, not to belief itself. And the contradiction fallibilism entails is among the fallibilist's beliefs. He believes each of his beliefs to be true and that some of them are not or might not be true.

Notice, by the way, that if we do not make the distinction between a belief and its justification or degree of being justified, then we would have to reject the laws of deductive logic, at least as applying to premises and conclusions that represent beliefs. In particular, we could no longer allow inferences of the form, "p. q. Therefore, p&q."

C.

Let us liken the set of a person's beliefs to the theorems of a formal system. Such an analogy is artificial and distorts the nature of belief. We will correct the distortion later. First, by exploring the analogy, we will see further into the question of fallibilism's inconsistency. Before setting up the analogy, however, we have to make some preliminary determinations of what "belief" means. We will not withdraw any of those determinations later; we will merely add to them.

Roughly, there are two approaches to the analysis of the concept of belief, one behavioral and one intellectual. The intellectual approach makes belief to be a proposition one affirms to oneself; hence, a belief is consciously thought and verbally formulated in a proposition. But one is not conscious of all his beliefs at once. Furthermore, it would seem that dumb beasts, that have no language in which to formulate propositions, also have beliefs. The dog that bounds door-ward, wagging its tail gladly, upon seeing its master pick up its leash, evidently believes that it is about to be taken for a walk. Hence the identification of belief with behavior and behavioral dispositions. The trouble with behavioral analyses is that beliefs are either true or false, while behavior and habits of behavior are neither true nor false. Also, there are habits that have nothing to do with belief, such as one's habit of putting on his left sock first.

Like many others after him, Peirce modified each view by combining them. We have already seen that he adopted the intellectual view, when he said that we *think* each of our beliefs to be true. In the same essay, only a few paragraphs later, he wrote that "belief is of the nature of a habit" (EP1:115), expressing the behavioral view. That was in 1877. A quarter-century later, he wrote of a "practical belief" as one that influences

what one does when acting "deliberately, bearing in mind his experiences, considering what he is doing, and exercising self-control" (5.538). For Peirce, deliberation and self-control involve consciousness: in 1905 and 1907, he associated consciousness with a level of self-control that requires deliberation (EP2:348,419). He concluded (in the same paragraph 5.538) that a practical belief is a "habit of deliberate behavior" – not any sort of habit, then, but one that combines thought and action. By contrast, a "theoretical belief" – such as that in the incommensurability of a square's diagonal and side – is more remotely related to conduct (5.539, cf. EP2:141); this passage should correct simplistic overstatements of Peirce's pragmatism. In 1895, Peirce wrote that "judgment is an act of consciousness in which we recognize a belief" – implying that a belief can exist unrecognized – "and a belief is an intelligent habit upon which we shall act when occasion presents itself" (EP2:19). As usual, there are other passages not easily reconciled with those preceding. The preceding, however, appear to be the ones most thoughtfully worked out.

The upshot is a two-fold indeterminateness of relationship between beliefs, *qua* propositional, and habits of action. On the one hand, a proposition, such as that sugar is sweet, can be translated into any number of habits of action, depending upon which other beliefs and desires and conditions it is combined with. My belief that sugar is sweet may determine me to add sugar to the bitter medicine I give my child but also to avoid adding sugar to my own tea, on the further belief that too much that is sweet is not good for me; and it may also lead me to call "Sugar" someone whom I find to be sweet, metaphorically. There is no end to such translations of a single belief into habits of action. On the other hand, an action or pattern of repeated actions may ground any number of alternative hypotheses about what a person or dumb brute believes. For a proposition, being verbally formulated, is always more specific than such evidence warrants (think, here, of Quine's argument for the indeterminateness of "radical translation"). Thus, a person who avoids plunging his hands into boiling water may be said to believe that doing so would cause him pain or a burn or injury – which? Even that person himself might not know which one is the exact formulation of the belief he acted on. The case is even worse for the dog bounding door-ward: is it a walk or exercise or the glorious smells of the great outdoors that the dog anticipates? The propositions in which we formulate canine beliefs are ours, not theirs.

What I propose, then, is that we think of a belief as a proposition on which a person would act in any of a wide range of circumstances and would formulate and affirm to himself in those or in other circumstances. The extension of this concept to dumb brutes is possible with qualifications; but we can leave that aside. In 1907, Peirce wrote that pragmatism "is scarce more than a corollary" of Bain's definition of belief as "that upon which a man is prepared to act" (EP2:399). If a belief is either true or false, then it follows that what one is prepared to act on is a proposition.

Putting aside the endless multiplicity of translations of belief into conduct, a person's beliefs, we might say, are all the propositions that he would affirm on reflection, without relying on any additional experience. So also, a formal system's theorems are all the formulae derivable in that system without the use of any additional premises. What is derivable is determined by the system's rules of inference (in which are included its axioms, if any). Now, one can know a formal system (that is, its rules) without knowing all its theorems; some work at derivation is required to arrive at any one theorem. And, typically, a formal system will have an infinite number of theorems. So also, there is no

time at which one can by reflection enumerate all that he believes. The set of a person's beliefs, like a set of theorems, is possibly infinite and is defined by what *can* be done (by that person reflecting), just as a system's theorems are defined as what *can* be derived by employing its rules.

We will now proceed to exploit this analogy.

D.

A formal system S consists of two sets of rules. One defines a set of permissible formulae. The other are the rules of inference that determine which formulae may be derived from others. To S we may add a semantics: a third set of rules by which the formulae, or some of them, may be interpreted as expressing propositions, either true or false. Let us assume that S has a semantics by which we can refer to some of its formulae as propositions (or as expressing propositions). The hope is that the rules of derivation are such that no false proposition can be derived from true propositions (that is, from a set of propositions of which none is false).

We shall read "$\Gamma \vdash p$" as a meta-statement about S (that is, not a statement that is itself a formula of S or expressible by any formula of S), to the effect that S permits the derivation of proposition p from the set of propositions Γ (all of these propositions being formulae of S or expressible by formulae of S according to S's semantics). If $\Gamma \vdash p$ and Γ is empty, then we may write "$\vdash p$", indicating that p is a theorem of S. We shall use "~p" both outside of S and within S to mean (assuming that S has a semantics) that p is false; it may be read "not-p." "$\vdash$ ~p" means that it is a theorem of S that p is false, while "$\sim \vdash p$" means that it is false that p is a theorem of S.

If S is a system of first-order predicate logic, then its formulae include predicates, individual constants denoting subjects, subject-variables, and quantifiers ("each" or "all" and "some" or "there exists"). The variables and quantifiers refer to a domain D of individuals (concrete or abstract). "$\forall x$" is the universal quantifier; we shall read "$\forall x(\varphi x)$" as meaning that every member of D has the property φ. And "$\exists x$" is the existential quantifier; we shall read "$\exists x(\varphi x)$" as meaning that there exists at least one member of D that has the property φ. $\forall x(\varphi x)$ and $\exists x(\sim \varphi x)$ are contradictories: not both can be true and not both can be false.

Now, S is consistent in the usual sense – we shall call it, "simply consistent" – if it has no theorems that are contradictory; i.e., there is no p such that $\vdash p$ and $\vdash$ ~p. Given that S possesses the usual rules of inference, including *reductio ad absurdum*, this means that S is simply consistent if and only if there is at least one proposition p expressible in S which is such that $\sim \vdash p$.

But a system consistent in this sense may be ω-inconsistent (the term and concept are due to Kurt Gödel). Suppose, for example, that for each member a_i of D, $\vdash \varphi a_i$ ("a_i" is the name of that member), but that $\vdash \exists x(\sim \varphi x)$. Then we perceive a contradiction even if none is derivable in S. None will be derivable in S if there is no way deriving "$\forall x(\varphi x)$ as a theorem of S, despite the fact that for each a_i, $\vdash \varphi a_i$. Suppose, for example, that D is infinite; then there is no finite time after which one can have derived all the propositions φa_i so as to draw the further conclusion, "$\forall x(\varphi x)$, contradicting $\exists x(\sim \varphi x)$. Also, whether D is infinite or not, S may lack a rule by which to generalize from a

complete enumeration of φa_i to the conclusion "$\forall x(\varphi x)$. S may lack means by which an enumeration can be identified as complete.

ω-inconsistency has been somewhat variously defined; but for our purposes it consists in a perception of a contradiction about a system that is not provable within the system. It depends on our being able to employ stronger forms of reasoning than the form of reasoning codified in S.

Note that if a system is ω-consistent, then it is simply consistent, but its being simply consistent does not entail that it is ω-consistent. Conversely, its being ω-inconsistent does not entail that it is inconsistent simply.

Bertrand Russell once gave this informal example of ω-inconsistency: it is a tautology that we want each of our desires to be fulfilled, but, among our desires, there is a desire for challenge, which depends on some desire not being fulfilled. So, we desire each desire to be fulfilled but we also desire that they not all be fulfilled. And is not that true? The very idea of being perfectly content is a dreadful sham. We are contradictory beings. And that is a very good thing.

Fallibilism is ω-inconsistent in precisely the same sense as is desire. We believe each of our beliefs to be true, but a fallibilist also believes that they are not all true. But while fallibilism is ω-inconsistent, it is not inconsistent simply – at least not within the analogy we have drawn between one's beliefs and a formal system S. To see this, let us define "├ p" for a system of beliefs, say, Jones' beliefs, parallel to our definition of the same term for a formal system S.

Just as, with respect to S, "├ p" means that p can be derived from S alone, without additional premises, let us say that "├ p" means that p would be affirmed by Jones were he to reflect on the matter, without recourse to further experience. In neither case does "├ p" mean that a derivation or reflection has actually been performed. However, one cannot know that ├ p, for some given p, without actually deriving it (or seeing in a general way how such propositions are derivable from S) or without Jones actually affirming it on reflection (or our seeing in a general way how such propositions probably would be affirmed by Jones once he reflected on the matter). (The two parenthetical clauses in the preceding sentence pertain to a meta-theoretical examination of S or of Jones, which is why, in the latter case, it is we, and not Jones only, who can make the determination.)

Now, if Jones can never finish enumerating his beliefs – thinking, redundantly, in each case, "That's true!" – because he has too many of them, or if he can never know when he has finished enumerating them, then, even if his mental capacities include that of drawing a general conclusion from a completed enumeration, he will never be able to arrive at the conclusion, "All of my beliefs are true." And therefore, even if he happens to be a fallibilist, he will never arrive at the contradiction, "All my beliefs are true, including the one that some of my beliefs are false." *We* (and he) can see, by taking an external or meta-theoretical overview of his system of beliefs that there is a contradiction among them; but it is not a contradiction that appears within the system itself. His beliefs are ω-inconsistent but they are simply consistent.

E.

Well, that's silly, isn't it? It depends on constructing an artificial idea of a "system of beliefs" – one in which there is the same rigid exclusion of meta-theoretical reflections as there is in a formal system. But belief is not like that. A fallibilist's meta-theoretical ruminations enter into his beliefs, producing an uncomfortable new one, that he is being inconsistent.

The way out of this imbroglio, I suggest, is to consider the actual part belief plays in practical life and in scientific inquiry, and to seek *in those terms* for a distinction something like that between simple consistency and ω-consistency. We shall say that the latter applies to one's beliefs considered abstractly as a set of propositions, and we shall limit the former the beliefs that, at any given moment, one actually acts upon. The abstraction, "the set of all one's beliefs," plays no part whatsoever in life or in science. We never act on all of our beliefs at once! And therefore a contradiction among them does not matter. All that matters, is that we do not attempt, at any one moment, to act on beliefs that are contradictory; for, ruling out no possible action, contradictory beliefs yield no decision.

But even that does not go far enough. Let us revisit an idea introduced earlier, that believing is justified when it is 99% certain. Obviously, the required percentage varies with the circumstances. If we are standing on a street corner in the rain trying to decide between Wong's Chop-suey Parlor and Hank's Hash House, it does not make much difference how sure I am that Wong's establishment is marginally better than Hank's. But suppose that we are physicians attending Jones who has an infection fatal if not treated; that the only antibiotics available are A and B; that A is sure to cure Jones if he is not allergic to it but is sure to kill him if he is allergic; and that B's efficacy is very much in doubt. Then it matters a great deal how certain we are that Jones is not allergic to A. Time and resources permitting, we will make careful tests to settle the question. Nothing less than near-certainty will satisfy us. Suppose, however, that there is no time for tests. Then we will act on the basis of what information we have, balancing the risk and cost of being wrong one way against the risk and cost of being wrong the other way.

Now, it is an interesting fact that where there is no time for more tests, a physician who has some reason to think that Jones is not allergic, might very well say, "I *believe* that Jones is not allergic; give him A," while, if there is time for more tests, he would *not* say that he believes that Jones is not allergic. He would suspend judgment until those further tests are made. If, before they are made, he is asked what he believes, he is likely to snap, "I have no belief about it!" The point is: what counts as belief is relative to the circumstances in which we are called upon to act.

And is not that a corollary of Bain's definition? If a belief is a proposition on the basis of which we are prepared to act, then one and the same proposition may be a belief in one circumstance and not a belief in others. The very idea that a person's beliefs are a set of propositions, to be considered in abstraction from changing exigencies, is mistaken.

One can seek to evade that conclusion by substituting propositions about probabilities for propositions that are less than absolutely certain. We might insist that the physician, instead of believing what he said he believed – "Jones is not allergic" – really believed, "The evidence available to me now makes it between 60% and 70%

likely that Jones is not allergic." And since decision was required, he thereupon concluded that the greater risk lay in not giving Jones drug A. But that is an artificial solution that saves our theory at the expense of misrepresenting the actual experience of decision-making, as reflected in how we ordinarily talk. Furthermore, it makes all of our beliefs to be propositions about probabilities. And it raises the prospect of an infinite regress: how probable is it that one's assessment of a probability is correct?

I must therefore disagree with Pascal Engel, who, in a valuable article in an earlier number of this journal (ENGEL 2005), argued for a pure sort of belief that he named "commitment," which is independent of practical exigencies and formed only in the pursuit of truth. I would rather adopt Peirce's view, that in that pursuit, our tentative conclusions are not beliefs. As theoretical conclusions, they are not beliefs; they are believed only if and when made the basis for practical decisions. As scientists themselves must decide which lines of inquiry to pursue and how to pursue them, Peirce overstated the case when he said that belief has "no place in science at all" (EP2:33; see MIGOTTI 2005 for a different way of looking at this issue).

In any instance, the beliefs on which one might act are those few that are relevant. If the question in non-dire circumstances is whether to test a proposition p that in dire circumstances we believe, then our fallibilist belief, that we can be mistaken, is relevant. On its basis, belief in p fades. "No," we say, "we might be mistaken about that. Better test it." There is no contradiction in that.[3]

References

CHISHOLM, Roderick (1952). "Fallibilism and Belief." *In*: WEINER, P.; YOUNG, F.H. (Eds.) *Studies in the Philosophy of Charles Sanders Peirce*. Cambridge, Massachusetts: Harvard University Press. p. 93-110.

ENGEL, Pascal (2005). "Belief as a Disposition to Act: Variations on a Pragmatist Theme." *Cognitio*, São Paulo, v.6, n.2, p.167-85, jul./dez. 2005.

MIGOTTI, Mark (2005). "The Key to Peirce's View of the Role of Belief in Scientific Inquiry." *Cognitio*, São Paulo, v.6, n.1, p.43-55, jan./jul. 2005

PEIRCE, Charles S. *Collected Papers of Charles Sanders Peirce*, v. 1-6. HARTSHORNE, C.; WEISS, P. (Eds.) Cambridge, Massachusetts: Harvard Universiy Press, 1931-35. [CP volume.paragraph]

_____. *The Essential Peirce*, v.1-2. Peirce Edition Project (Eds.) Bloomington, Indiana: Indiana University Press, 1992-98. [EP, volume: page]

---

[3]   The only discussion of this problem that I have found in the literature is by Roderick CHISHOLM (1952), who focuses on the special case of common-sense beliefs. His solution is that they can be doubted when considered together but not when taken up individually. That solution subverts deductive logic and makes sense only if we identify belief with its probability. It falls short of our conclusion, that no proposition is a belief except contextually, when it is a basis for action.