

This Gödel is killing me: a rejoinder (*)

J.R. Lucas

Fellow of [Merton College](#), Oxford
Fellow of the [British Academy](#)

Mr. Hutton claims (1) that we cannot use Gödel's theorems to prove minds different from machines, because we are not entitled to assert our own consistency. Maybe not. But not for the reasons Mr. Hutton urges. They seem to me to have very little to do with the argument I originally put forward, although something to do with the logical analysis of our everyday concepts and the relation between formal and informal modes of reasoning.

Mr. Hutton's arguments are statistical. He seeks first to show that there is a non-zero probability of a mind's embracing mutually inconsistent propositions; secondly that such a possibility might actually be realised in the person of a gambling logician; thirdly he seems to argue that any claim to be consistent must be a claim to infallibility which is insensitive to counter-arguments to the point of irrationality. I do not deny that many people are inconsistent. In my original article I conceded that women and politicians were popularly supposed to be, and allowed that male non-politicians might be too. If Mr. Hutton cares to cite gambling logicians as well, I shall not dispute with him. But I shall remain sceptical of his first argument, in which he considers an n-dimensional discrete mental life-space, consisting of n propositions (or later axioms) and their negations, and assigns probabilities to a mind's accepting a subset of this space, and calculates that the probability of a mind's accepting an inconsistent set of propositions is very high. Quite apart from the general artificiality of this model of the mind, it seems to have one decisive defect even from the technical point of view. Mr. Hutton assigns probabilities to a mind's accepting each proposition *independently* of what that mind has done hitherto, whereas most people are guided, in deciding whether to accept or reject a particular proposition, by what they have already taken to be true. The statistical considerations Mr. Hutton adduces have no bearing on the informal reasoning I was appealing to.

Mr. Hutton's third argument relies on Bayes' Theorem. I have [145] grave doubts about the use of Bayes' Theorem for singular propositions, such as my being consistent (C). It seems to me that Bayes' Theorem, and the underlying axiom of Conditional Probability, applies only to propositional functions ranging over certain universes of discourse, and that when we are assessing singular propositions on the basis of arguments and counter-arguments, we need, if anything, some sort of Confirmation Theory rather than the classical Probability Calculus. (2) Not that the classical Probability Calculus is inconsistent, but rather that it is inapplicable. The axioms of classical Probability Calculus are consistent and apply very well to additive sets, and under suitable conditions can be used to calculate the probabilities of propositions. But not all conditions are suitable, and often we use the word 'probability' in a way which is not adequately rendered by Kolmogorov's axioms. More serious than doubts about Mr. Hutton's handling of probabilities - about which many different views are canvassed should be doubts about his assumption that a person who knows something cannot allow that he might be wrong. Locke held this (3) and later Ayer, but it led them to limit knowledge to immediate sense-experience and analytic truth. Normally, however, the word 'know' is used in a much wider sense. I can properly say that I know Mr. Hutton exists, although I have never seen him, and there would be no logical impossibility in doubting it, and I could be persuaded, say by a

sufficiently authoritative statement by the editor, that he was merely the *nom-de-plume* of Professor Bar Hillel. Nevertheless, I can properly say not merely that I believe Mr. Hutton exists but that I know it. For in addition to actually holding that belief, I have good reasons for so doing: not incontrovertible reasons nor conclusive reasons, but, in the circumstances and in the context of my statement, adequate ones. To claim to know something is not to claim infallibility but only to have adequate backing for what is asserted. It is a responsible guarantee, but not a copper-bottomed one. If I use the formula 'I know that...' I utter a negotiable instrument in the republic of letters. You can pass it on to third parties on the strength of my warranty, and if you are challenged you can say 'I have Lucas's word for it'. Should he then have recourse to me, I am obliged to honour by bond, and give the reasons which backed it; and if my account turns out to be empty, or if I have only slight reasons, or if I have overlooked obvious objections, then you are entitled to hammer me for issuing promissory notes irresponsibly. But I am not required to have more than a reasonable certitude. If I claim to know the time of the next train, it is enough that I have looked it up in the timetable: I do not have to have considered the possibility of its having been derailed on the way to the station. Under some conditions, when there are wars or rumours [146] of wars, we may have to hedge all our bets, and rather than claim to know the time of the next meeting of the Knesset, express only a tentative belief that it may meet next month, provided the Syrians do not invade, nor any other comparable contingency occurs. But such conditions are exceptional. In normal times, although we could always add a *Deo Volente* when making any sort of prediction, we do not do so, nor are our claims to knowledge any the worse for the omission.

To assert one's own consistency is not exactly like predicting the time of the next train or next meeting of the Knesset. It is more like knowing that one is conscious, or knowing that other minds exist, or knowing that the future will be like the past. These are general presuppositions on which much of our thinking is based. They cannot be proved by deductive or inductive reasoning, and they can be argued against, and some philosophers have schooled themselves to regard them as open questions. But although they can be argued against, they can be argued for, and few people, except when they are being paid to as professional philosophers, seriously doubt them. In the case of the mind's consistency, I do not deny that there are arguments against my claim. In addition to counter-examples among our acquaintanceship, I cited naive set theory as a warning against too easy an assumption that what seemed all right intuitively was all right really. Nevertheless, the arguments on the other side are weightier still. One, which Mr. Hutton touches on but does not meet, is that simple consistency is, in the relevant systems, equivalent to absolute consistency, and it is highly implausible to make out that we are absolutely inconsistent. We say it occasionally as a reproach to a person "If you would say that, you would say anything", but it is not in general true of most people. If a mind could be adequately represented by an inconsistent system, then it would be prepared to affirm every proposition. But most people will not do that. Therefore they cannot be represented by inconsistent systems.

Another argument, of a more Kantian flavour, is that we must assume our own consistency, if thought is to be possible at all. ⁴ It is, perhaps like the uniformity of nature, not something to be established at the end of a careful chain of argument, but rather a necessary assumption we must make if we are to start on any thinking at all. (4) In a somewhat similar vein I suggested that it should be seen as an act of faith, a decision or a commitment, nor arbitrary although not admitting of formal proof. Some men similarly decide to treat other human beings as centres of consciousness and autonomous agents.

None of these arguments lends any colour to the contention Mr. Hutton wishes on me that I know that I am consistent in [147] some dogmatic way which precludes all possibility of further debate. He does not deal with the arguments I actually adduce, but quotes instead two other passages dealing with different topics (one about the infinite regress of consciousness,

the other about Tarski's theorem) to impute to me a dogmatic assertiveness quite foreign to my argument. My only purpose was to counter an objection by a mechanist who claimed he had a model of the mind and sought to deny purchase to the Gödelian argument by claiming that his model either was not, or at least could not be taken to be, consistent. Against this I argue that an inconsistent system cannot be worthy of serious consideration as a model of the mind, and that unless the mechanist claims consistency for his model - in which case the Gödelian argument goes through - it is not a plausible candidate at all. Certainly, if once an inconsistency is spotted in a system, it ceases to be of much interest to us; and if I could spot an inconsistency in the system which was claimed by the mechanist to represent my mind, I should at once point out that $2 + 2 = 5$ was a theorem of the system but no theorem of mine. No doubt mechanism is an implausible thesis on many other counts than those urged by me. My object was to take on the mechanist on his own ground, and show that, all other difficulties apart, there were serious logical difficulties in putting forward a mechanical model of the mind. The main thrust of the argument was against models that were at least consistent, but the possibility of inconsistent models had been canvassed; they were, however, implausible models of a rational mind, and, in spite of Mr. Hutton's arguments, seem to be so still.

Notes

(*) First published in *Philosophia*, 1976, 1, pp. 145-148. For the copyright of the papers of J.R. Lucas see <http://users.ox.ac.uk/~jrlucas/> [back](#)

(1) Anthony Hutton, [This Gödel is Killing Me](#), *Philosophia*, this issue. [back](#)

(2) I have developed this point more fully in J.R. Lucas, *The Concept of Probability*, Oxford, 1970, esp. pp. 101-103, 146-162. [back](#)

(3) See e.g., John Locke, *Essay Concerning Human Understanding*, Bk IV, Ch. 6, section 16; Ch. I 1. section 9. [back](#)

(4) J.R. Lucas, [Minds, Machines and Gödel](#), *Philosophy*, XXXVI, 1961, p.124; reprinted in *The Modeling of Mind*, Kenneth M. Sayre and Frederick J. Crosson, eds., Notre Dame Press, 1963, p.268; and *Minds and Machines*, ed. Alan Ross Anderson, Prentice-Hall, 1954, p.56.). Or J.R. Lucas. *The Freedom of the Will*, Oxford, 1970, pp. 161-163. [back](#)