What logic should we think with?

Logic ought to guide our thinking. It is better, more rational, more intelligent to think logically than to think illogically. Illogical thought leads to bad judgement and error. In any case, if logic had no role to play as a guide to thought, why should we bother with it?

The somewhat naïve opinions of the previous paragraph are subject to attack from many sides. It may be objected that an activity does not count as thinking at all unless it is at least minimally logical, so logic is constitutive of thought rather than a guide to it. Or it may be objected that whereas logic describes a system of timeless relations between propositions, thinking is a dynamic process involving revisions, and so could not use a merely static guide. Or again the objection may be that there is no such thing as logic, only a whole variety of different logics, not all of which could possibly be good guides.

I aim to disarm the last two objections to the initial idea that logic should be a guide to thought.

1. Logic and belief revision

How could logic guide our thought? Anything we can believe has an infinite number of logical consequences, but no sensible guide would tell us to take steps to believe all the logical consequences of anything we believe, for there is not enough time to obey and most of the added beliefs would be trivial.

There is something to the idea that we should consider our beliefs not just one by one but in larger groups, to see if we can logically extract some useful further information. We might try to implement this idea by the following prescription:

1. If some set of propositions X logically entails a proposition A and one believes each member of X then one should also believe A.

One problem has already been noted : (**Error! Reference source not found.**) would require us to believe too much useless stuff (we would, for example, be endlessly forming conjunctions of our beliefs, to no good purpose). So an account of how logic might guide thought must somehow concern only some subset of logical space, and must balance processing costs against cognitive gains.

The problem with (**Error! Reference source not found.**) is not confined to the excessive number of propositions it would have us believe. A speedy being who, unlike us, could form one new belief in one second, another in half a second, a third in a quarter of a second, and so on, could form an infinity of beliefs in a couple of seconds. Prescription (**Error! Reference source not found.**) does not give good advice even to such a speedy creature, for it takes no account of dynamic updating. Suppose such a creature set out to Gordon Square on Thursday 12th October 2000 believing that a logic lecture is to be held there on that day. Finding on his arrival that there is no lecture, he is not required, by logic or anything else, to believe that there both is and is not a lecture, despite the fact that A, B † A & B; that is, despite the fact that the contradiction is entailed by things he believes. Prescription (**Error! Reference source not found.**) makes it sound as if all a rational creature does is add beliefs; there is no account of how they are taken away.

Logical entailment is one crucial logical relation, and its natural application as a guide is additive. But (**Error! Reference source not found.**) provides nothing to control increase. We do better, I think, to concentrate on the closely related logical relation of inconsistency. In its immediate application, it decreases rather than increases beliefs: discovery of an inconsistency

in our beliefs should make us abandon a belief. A first attempt to exploit inconsistency in finding a role for logic as a guide might be this:

2. One whose logical commitments are logically inconsistent should revise his beliefs in such a way that the commitments are consistent.

The main problem is that the inconsistency might be unrecognized, in which case the thinker would have no good reason to revise one of his commitments. Here is an example which illustrates this:

On Monday, James tells you that Jane was at the party and you believe him. On Wednesday, George tells you that Jane was not at the party and you believe him. You do not revise your previously formed belief, because it does not come to mind. If you were asked on Friday whether or not Jane was at the party, it's a toss up whether you would say Yes or No or something else. Under these circumstances, you've done nothing irrational. The irrational thing would have been to engage in belief revision with no reason before your mind.

It may be objected that we have a standing reason to scan all the time for inconsistency. But this is not a reason we recognize. Constant scanning would take up time and energy; we need a specific reason to use up these resources, and the mere abstract possibility of an inconsistency is not enough. Moreover, the prescription is too general: it does not give any clue about where to look for inconsistency, and this increases the probability that the effort would be wasted.

At least part of the problem is addressed by the following:

3. One who realizes that his logical commitments are logically inconsistent should revise his beliefs in such a way that the commitments are consistent.

But suppose the inconsistency is completely unimportant to you? Reverting to the displayed example, suppose you have no interest in whether or not Jane was at the party? Why should you invest effort in adjusting your beliefs, even if you recognize their inconsistency? That you should not becomes plainer when we consider the options in more detail:

(a) Resolve the contradiction by seeking out further information.

This may be impossible or absurdly costly. James and George may both now be dead, you may have no access to any other Jane-related informants, and so on. It would be crazy to say that logic requires you to hire a private detective to help gather evidence which might resolve this trivial matter.

(b) Abandon one or other of the beliefs without seeking further information.

But how to choose? As James and George are equally reliable, there is no reason for either choice.

(c) Abandon both beliefs.

This sounds appealing, but it may be a bad strategy. Suppose you learn that if Jane was at the party something terrible will happen to you unless you take a low-cost preventative measure, P. If you have simply abandoned your belief that Jane was at the party, this new information gives you no reason to do P. But intuitively you do have a reason.

The problem with (**Error! Reference source not found.**) was that it paid no attention to the relation between the costs and the benefits of revision. We can at least gesture towards the needed comparison as follows:

4. One who realizes that his logical commitments are logically inconsistent has a defeasible reason to try to revise his beliefs in such a way that the commitments are consistent.

I do not hold this up as a final statement of a way in which logic could act as a guide. (For one thing, a more finished account should talk in terms of degrees of belief and commitment.) However, it should encourage optimism about the prospects for a substantive and intuitively satisfactory relationship between the static relations of logic and the dynamics of thought. A defeasible reason is one that can be overridden by others, and the range of potentially defeating

reasons are intended to include the ratio of the importance of the topic to the amount of effort that would be required to arrive at a well grounded resolution. (Whether "ratio" is to be taken literally as a number is one of the many issues that would arise in a detailed study of this issue.)

Prescription (**Error! Reference source not found.**) is not intended to exhaust the possible ways in which logic could properly guide thought. It is consistent with the opinion that a logical principle should also serve as a guide in the sense of something we (perhaps implicitly, thanks to good training) consult when we think.

2. Which logic should we choose as our guide?

When people speak of logic without qualification, they probably mean classical logic: the logic of first order quantification with identity. This is a recent phenomenon, and we should not forget the enormous variety of alternatives to classical logic that have been proposed, of which the following list gives a sample of points of divergence from classical logic:

- Double negation:
  Classical logic accepts the inference from "not not A" to "A"; intuitionist logic does not.
- Existential import
  Classical logic rejects the inference from "All Fs are G" to "Some Fs are G"; Aristotelian logic accepts it.
- Disjunctive syllogism:
  Classical logic accepts the inference from "A or B" together with "not A" to "B"; relevance and paraconsistent logics do not.
- Modus ponens:
  Classical logic accepts the inference from "if A then B" together with "A" to "B"; some fuzzy logics do not.
- Ex Contradictione Quodlibet:
  Classical logic accepts the inference from "A and not A" to "B", where "B" is an arbitrary proposition; paraconsistent logic does not.
- De Morgan
  Quantum logic rejects de Morgan's laws.
- Instantiation:
  Classical logic accepts the inference from "Everything is F" to "α is F"; free logic does not.
- Negation
  Classical logic holds that negation toggles between truth and falsehood; option negation denies this.
- Conditionalizing disjunctions:
  Classical logic accepts the inference from "A or B" to "if not A then B"; probabilistic logic does not.

Logics can differ about what they count as an inconsistency. For example, within some free logics the beliefs that *everything is self-identical* and *Vulcan is not self-identical* are consistent, but they are classically inconsistent. So the objection that we cannot simply hold that "logic" should be our guide is correct. We have to choose the right logic. But how could this choice be made? Will not making it inevitably involve reasoning, in which case some logic is already presupposed? The remainder of this paper addresses this methodological issue. I introduce it by considering how one might argue for or against the last three non-classical logics on my sample list.

Instantiation:
Classical logic adopts the following rule of instantiation:

  If $X \mapsto \forall v \varphi v$ then $X \mapsto \varphi\alpha$

If some set of sentences entails a universal generalization, then that same set of sentences entails what results by deleting the quantifier in the generalization and putting a name in place of every occurrence of the quantifier's variable. A natural application to English suggests we should regard the following as valid:

  5. Everything is perishable, therefore Socrates is perishable.

But intuitively this is not valid: it could be that everything is perishable and Socrates is not (if he did not exist).

This is straightforwardly accommodated by the free logical instantiation rule:

  If $X \mapsto \forall v \varphi v$ then $\{X, \exists v(v=\alpha\} \mapsto \varphi\alpha$

We have to supplement the premises which sustain the generalization with one which affirms the existence of the object we will use in the instantiation. This supplementation does not occur in (**Error! Reference source not found.**), so (**Error! Reference source not found.**) is not validated by the free logical instantiation rule.

We have here an argument, eminently defeasible of course, for preferring some version of free logic to classical logic: free logic respects intuitions about validity which classical logic does not.[1]

The argument is defeasible because it provides no reason for thinking that only free logic can deliver appropriate judgements of validity. Perhaps other adjustments to classical logic can conform to the alleged datum (the invalidity of (**Error! Reference source not found.**)), and this shows that the argument, even if good against classical logic, does not point specifically towards free logic.

It is also defeasible because the claim on which it rests, the invalidity of (**Error! Reference source not found.**), can itself be challenged. Perhaps when we assess validity we should hold the domain constant. When we ask after the validity of an argument we start with the objects actually involved in the premises, whether as objects quantified over or objects referred to; the worlds relevant to validity are those containing just these objects. On this ruling, only worlds containing Socrates are relevant to the validity of (**Error! Reference source not found.**), and the case envisaged, in which Socrates does not exist, is irrelevant. So we should reject the alleged datum.

This suggestion does not accord with classical semantics, so it would in any case involve moving away from classical logic, even if not in the direction of free logic. But the suggestion about validity does not accord with our intuitions or our best judgements. Consider the following argument:

  6. Everything is a duck; so everything flies.

Suppose that all actual ducks are essentially fliers, but that if evolution in New Zealand had taken a slightly different course their range of flightless birds would have included flightless ducks. Holding the domain constant would rule that (**Error! Reference source not found.**) is valid: every world containing just the ducks our world contains verifies the conclusion if it verifies the premise. But intuitively the possibility of flightless ducks shows that it could be that everything is a duck without everything flying, and this intuitively establishes the invalidity of the argument.

---

[1] I first heard the argument from Yannis Stephanou, though he does not himself endorse it.

Classical logic could accommodate the validity of (**Error! Reference source not found.**) by resisting the formalization of the premise as no more complex than "$\forall x Fx$". Why not treat "Everything is physical" as having the same form as "Every ring is circular"? Then (**Error! Reference source not found.**) would be formalized by the invalid

$$\forall x(Fx{\rightarrow}Gx) \mapsto G\alpha$$

which shows that the original argument was wrong not in its datum (the invalidity of (**Error! Reference source not found.**)) but in its assumption that this could not be accounted for classically: (**Error! Reference source not found.**) is invalid because it lacks the premise "Socrates is a thing" ($F\alpha$). (A similar result would be attained within an essentially classical framework by formalizing with binary quantifiers.)

In this discussion free use has been made of intuitions (beliefs held with justification but without ground) and of principles of reasoning. There was no sense that we could not start reasoning until we had resolved the question of which logic to use as our guide.

Option negation:
The capacity for deliberative choice is at least no less developmentally basic than the capacity for language. Any language user must be capable of deliberative choice, and arguably some non-language-users (for example, apes) have the latter capacity. It is therefore not crazy to look to the preconditions for deliberative choice for guidance about the character of language. The suggestion I wish to consider is that the exercise of choice requires a concept of negation which has had a hand in moulding the behaviour of current words for negation.

A chooser must take herself to be confronted with a range of options. We think of the options as exclusive, and the suggestion is that this exclusivity generates the distinctive concept of option negation. I see my options for action as A and B. To select A is to deselect B: NOT-B expresses this deselection. In the context, A and NOT-B are equivalent, and so are NOT-A and B. If we think of the options as marking the truth or falsity of some statement, option negation will behave classically: deselecting truth amounts to selecting falsehood, that is, NOT-true amounts to false; and NOT-false amounts to true.

Within the structure of options, option negation is a more general notion than classical negation, and can explain otherwise puzzling features of negation. For example, not only are the following never normally heard as inconsistent, it is impossible to hear them as inconsistent:
    7. It's not a car, it's a Volkswagen.
    8. No, I have not *pak edd zee soo eet cass ez*. I have packed the suitcases.[2]
Presumably the advertising agency thought there would be something striking about (**Error! Reference source not found.**), but they certainly did not suppose the audience would infer that a Volkswagen is not a car, which would be a classical consequence. One could imagine (**Error! Reference source not found.**) uttered with some irritation as a response to "Af you pak edd zee soo eet cass ez?" by one who has tried to help the speaker with her pronunciation. Again, it is classically inconsistent, but it is impossible to hear it thus. This is hard to explain classically, since any "pragmatic" account should presumably start with perceived literal inconsistency, and that is just what, I think, cannot be perceived.

---

[2] Examples of this kind abound in Horn, L. R. (1989) *A Natural History of Negation*. Chicago: University of Chicago Press, who uses them to ground a distinct notion of negation which he calls "metalinguistic negation".

Within the option negation framework, all is straightforward. In (**Error! Reference source not found.**) the options are to call it a car or to call it a Volkswagen, and the former is deselected. In the (**Error! Reference source not found.**) the options are to say "I af pak edd zee soo eet cass ez" or to say "I have packed the suitcases" and the former is deselected. All interpretation involves bringing to bear something more than what is usually counted as semantic knowledge, for example a hearer must at a minimum know what language the speaker is speaking. In the case of option negation the hearer must latch on to the relevant options, and an account must be given of how this is possible. The problem seems much more tractable within this framework, however, for a basic notion in dialogue is what the right thing to say is, where the right thing does not equate to a true thing, and applying this notion to both (**Error! Reference source not found.**) and (**Error! Reference source not found.**) gives an appropriate result: the right thing to say is that it's a Volkswagen, not that it's a car, and the right thing to say is "I have packed the suitcases", not "I af pak edd zee soo eet cass ez". This seems much more straightforward, and amenable to systematic theory, than an account which begins by attributing to the hearer a realization of the literal inconsistency of (**Error! Reference source not found.**) and (**Error! Reference source not found.**).

It is a striking feature of our conception of options for action that the discovery that two options we have been considering can be combined is represented as a third option: A, or B, or both. (The discovery of this third option seems often to be made by the more intelligent of parties to a discussion of what to do.) Option negation now manifestly differs radically from classical negation, for NOT-both amounts to A or B; and NOT-A amounts to B or both. "Both" differs from "and", for "both(A,B)" entails "NOT-A" and also "NOT-B". Likewise "neither" is a further option, so a common choice situation is A, B, both or neither. Such a four way partition could be used to model a system of truth values allowing for both gluts and gaps, as well as plain truth and falsehood, so the options for a statement are: true, false, both or neither. "Both" entails "NOT-true" and also "NOT-false".

This may seem mere idling: anyone can invent a strange particle which behaves in odd ways. In defence of option negation, I plead its basis in the notion of choice, and its capacity to explain oddities about our actual uses of negative particles. Here is a further oddity. Berkeley accused Locke of "manifest contradiction" when, in speaking of the abstract idea of a triangle, he said that

> it must be neither oblique nor rectangle, neither equilateral, equicrural, nor scalenon; but all and none of these at once. (*Essay* 4.7.9)

I suspect I am not alone in thinking that there is a natural and consistent interpretation of Locke's words "all and none of these at once". We cannot consistently apply both "equicrural" and "scalenon", but I wish to focus just on whether "all and none" is automatically inconsistent. Within the option negation framework, it is not. Taking "all" as an extension of "both", it actually entails NOT-*o*, for each other option *o*, and "none" could be seen as summarizing these entailments.[3]

What counts as inconsistent within an option negation framework differs from what counts as inconsistent within the classical framework, so if logic is to be a guide in anything like the envisaged way we have to make a choice. The point of the present discussion is not to make that choice, but only to point out that one cannot rule out in advance the possible relevance of empirical data about how we have developed as agents. It is supposed to be an empirical and

---

[3]At a conference on the history of logic at King's College London in November 2000, Graham Priest cited Plotinus as an example of an early dialetheist: "The One is all things and no one of them." If Plotinus was operating in the option negation framework, we need not see him as affirming anything inconsistent.

historical fact that our actual use of negation grew out of option negation. This is in principle susceptible of aposteriori confirmation or disconfirmation. This suggests that it is possible to have evidence of an aposteriori kind for or against a logical principle, and that this evidence can be what Field has called primary: it is evidence which in principle could outweigh non-empirical evidence.[4]

Conditionalizing disjunctions:
Classical logic accepts the inference from "A or B" to "if not A then B"; probabilistic logic does not. So classically, but not probabilistically, the combination "(A or B) but not: (if not A then B)" is inconsistent. We need to choose between these logics.

The discussion is not (and could not be) supposed to impugn the validity of the pattern:
$$A \vee B; \text{ so } \neg A \rightarrow B.$$
But if a logic is to apply to our thought, it must do so via our natural expression of thought, our natural language. Strictly speaking, then, the first paragraph of this section should claim that, as most obviously applied to natural language, classical logic accepts the inference from "A or B" to "if not A then B".

A much discussed instance of the principle is the following:
9.   Either the gardener did it or the butler did it; so if the gardener didn't do it, the butler did.[5]

The following seems a reasonable system of belief:
   The most likely hypothesis is that the gardener did it and got the butler to act as his accomplice, but the butler would never have acted alone: he is too decent a chap, just suffers from weak will, and the gardener can be very persuasive.

One who has these beliefs is committed to thinking it highly likely that either the gardener did it or the butler did it. But it seems that he could hardly think it at all likely that if the gardener didn't do it, then the butler did. On the contrary, he must think that if the gardener didn't do it, it's very unlikely that the butler did: the butler would not have acted alone. So it seems to be reasonable to assign a high probability to the premise and a low probability to the conclusion. Yet if an argument is valid in the classical sense, the conclusion has to be true if the premise is, and so it would not be reasonable to assign the premise a much higher probability than the conclusion.

The same point can be illustrated with numerical assignments. A probability function which makes the following assignments seems perfectly reasonable:
   Pr(The gardener did it) = 0.9; hence
   Pr(The gardener did it or the butler did it) ≥ 0.9
   Pr(If the gardener did not do it, the butler did) = 0.1.

The last probability is a consequence of one's supposition that it is highly unlikely that the butler acted alone. An argument with a premise to which it is reasonable to assign high probability and a conclusion to which it is reasonable (in the same informational state) to assign low probability is not valid.

Some other examples of the same pattern of argument are more likely to strike one intuitively as invalid:

---

[4] Field, Hartry (1996). "The apriority of logic." *Proceedings of the Aristotelian Society* 96: 359–379.

[5] Cf. Stalnaker, Robert (1975). "Indicative conditionals." *Philosophia* 5: 269–86. Reprinted in Jackson, Frank (ed.), *Conditionals,* Oxford: Oxford University Press, pp. 136–154.

Either I'll catch my train or I'll be on time for work; so if I miss my train I'll be on time for work.[6]

Either he'll throw an even number or he'll throw a 6; so if he doesn't thrown an even number, he'll throw a 6.

It may be objected that the argument proves too much. Suppose modus ponens is valid. Even so, one could reasonably assign high probability to A and to if A then B without automatically assigning high probability to B. For example, I believe that there are some unsold tickets and that if there are some unsold tickets I will be able to buy one at the door. I am highly confident in both beliefs, which entail that I will be able to buy one at the door. Yet when I arrive I find that I cannot buy one at the door: they have all been sold in an unexpected late afternoon rush. I then assign zero probability to the proposition that I will be able to buy one at the door. Should we conclude that modus ponens is not valid? Obviously not. But this case differs from the one relating to the gardener and the butler (**Error! Reference source not found.**), for in the present case the high probability assigned to the premises does or should change when the new information revises the probability of the conclusion down to zero. In (**Error! Reference source not found.**), by contrast, the probability assignments are rationally cotenable in a single informational state.

I conclude that (**Error! Reference source not found.**) is indeed invalid and hence that either classical logic must be revised or it must not profess to be able to formalize English conditionals as material conditionals. As in the previous cases in which an alternative to classical logic has been considered, the point has been not to demonstrate that the alternative is or is not correct, but to prepare for methodological reflections on how a logic could be chosen (without some logic having already been chosen in order that there is a logic to guide the choice). In the present case, intuitions about reasonable degrees of belief were used to evaluate a logical principle. The intuitions are subject to empirically or apriori based disconfirmation, for example Dutch Book arguments.

3. Methodological conclusion

The arguments sketched for or against the three non-classical logics (free logic, option negation logic, and probabilistic logic) did not appeal to some unchallengeable core of logic that must be taken for granted in all thought or discourse. The arguments appealed to considerations of completely different kinds. In the case of free logic, an allegedly invalid argument was produced, and was said to be a counterexample to classical logic. The only support offered was intuition. The proposed verdict on the argument did not depend upon challenging the intuition of invalidity, but rather on challenging its status as a counterexample: with appropriate formalization, its invalidity could be accommodated within the classical framework. In this discussion, everything was apriori. By contrast, historical and psychological considerations were supposed to be available to support option negation. Apriori reasoning would be involved as well, if the resulting logic were to be developed, but an essential empirical component would remain. In the final case, of conditionalizing disjunctions, logical validity was tested by its impact on subjective probabilities. What was drawn from the area, for example the general principle that a rational person could not, relative to a single state of information, assign greater uncertainty to the conclusion of a valid argument than to the sum of its premises, was largely apriori, but in contrast to the first case the considerations were indirect. We were not asked to intuit the validity status of some argument: rather, reasons were provided for making a possibly counterintuitive judgement of its status. The probability theory which was relied upon, though

---

[6] Cf. Edgington, Dorothy (1995), "On conditionals" *Mind* 104: 235–329.

apriori, is not merely intuitive, for it is well known that intuitions about probabilities are often demonstrably defective.

The mere fact that the considerations were different in each case undermines the thought that there is an unchallengeable core from which we should work outwards in trying to determine which is the correct logic. This diversity is apparent with other alternatives to classical logic. For example, the case for intuitionistic logic is typically grounded on general considerations about the nature of meaning; the fate of bivalence has been linked with that of metaphysical realism; quantum phenomena have been said to undermine de Morgan's Laws; and various revisions of classical logic have been suggested in order to accommodate specific problems (for example, sorites reasoning).

We ought to have encountered an is/ought problem. In this kind of context, we can adduce how or what we in fact think, but how can we simply *adduce* how or what we ought to think, for that is the very question at issue? Will we not have found ourselves moving invalidly from facts about what we take to be so to facts about what we ought to take to be so? This did not appear to be a pressing issue in the examples sketched, and this fact requires some explanation.

An appealing general view in epistemology is that our beliefs are "innocent unless proved guilty". Less metaphorically, the "innocence principle" is that we are justified in our beliefs, unless we have reasons for suspicion. (So "*proved* guilty" is too strong: any decent criticism needs to be examined, and if left unexamined defeats justification.) To say that we are justified is not to say that we have further beliefs that provide justifying reasons, but is only to say that we are justified. Holding the beliefs which the innocence principle says are justified is not subject to coherent specific criticisms (as things stand right now) for any such criticism would automatically count as a suspicion, and criticized or suspect beliefs are not held by the principle to be justified.

We could extend this to our logic: our inferences, likewise, are innocent unless proved guilty, or, more accurately, unless there is some ground for suspecting them. We are justified in reasoning as we do, unless we discover some reason for thinking that these ways are not good. If correct, this would explain why we have encountered no is/ought problem.

Those of our actual inferential practices which we have no reason to suspect are ones in which we are justified in engaging. This means that in appealing to some inferential practice, even in the course of throwing doubt upon another, we are appealing to an inferential pattern we are justified in using, and this takes us straight over the supposed gap between "is" and "ought": if this is in fact how we reason, and we have no ground for suspicion of this reasoning, then this is how we should reason. The general idea is the same as in non-foundationalism about empirical justification: nothing is sacrosanct, but we have to start somewhere and there is no thinkable alternative to starting with the beliefs we actually have at some particular moment and which (at that moment) we have nothing against. The logical boat, as much as the empirical one, must be reconstructed at sea, plank by plank, with no plank privileged and with no pre-ordained right starting-point. On this picture, any logical principle can in theory be criticised, and we need not first establish the correctness of the logical principles lying behind the criticism, though these themselves may be targets of criticisms in due course.

The innocence principle has been criticised. One must certainly accept that it does not tell the whole story about justification: it does not say how it is transmitted or defeated or how revisions should occur. But it has also been claimed that the principle does not state a sufficient

condition for justification: a justified belief needs to have a suitable aetiology.[7] There are at least four ways in which this could be absent:

10. The belief is reached on the basis of bad reasons (ones which do not genuinely support the belief).
11. The belief is of a kind which requires other beliefs as reasons if it is to be justifiably believed, and no such reasons are held.
12. The belief is a product of a cognitive mechanism which is operating defectively.
13. The belief is a product of a cognitive mechanism which is operating under conditions for which it is not fitted.

Illustrations:

Of (**Error! Reference source not found.**): One person believes that there is a £50 note in his pocket through a combination of perception and memory: he has just looked to check. His belief is justified. By contrast, someone who reaches the same belief merely through wishful thinking is not justified, even if she has no other belief which would cast suspicion on this one.

Of (**Error! Reference source not found.**): A complex mathematical statement is properly believed only on the basis of a proof. This assertion is potentially directly inconsistent with innocence, but could be supported by the following contrast. Two people get to a certain point in a proof of some complex mathematical claim, C, from which C follows immediately. One person sees this immediate inference and on the strength of it comes to believe C. The other does not see the inference, but believes C anyway, saying to himself "Well, this simply has to be true". It seems that although the first person is justified the second is not, even if he has no grounds to suspect the conclusion.[8]

Of (**Error! Reference source not found.**): Jaundice is supposed to make white things look yellow. Someone suffering from this disease who believes that something is yellow is arguably not justified in that belief. Although cognitive mechanisms sometimes make manifest that they are not working properly (e.g. if some part of the body goes numb, the sense of touch is manifestly lost in that part), this is not always so.

Of (**Error! Reference source not found.**): If one's head is under water, sounds are distorted. A belief formed on the basis of hearing in this circumstance is arguably not justified. Although cognitive mechanisms sometimes make it manifest that they are in an environment for which they are not fitted (e.g. if it is dark, our visual system manifestly does not work) this is not always so; we might not appreciate the distorting influence on hearing of being under water.

On this view, innocence is an oversimplified version of the following truth: a belief with a satisfactory aetiology is justified, provided there is no ground for suspicion.

For present purposes, we can allow these illustrations of the alleged failure of innocence to pass unchallenged, for, I shall suggest, the justification of inferential practices, including tendencies to regard certain collections of propositions as inconsistent, has special features which give a different slant to these kinds of criticism.

The contents of some beliefs, like the one about the £50 note in the illustration of (**Error! Reference source not found.**), demand a specific kind of aetiology, one which is ultimately

---

[7] Pollock, John L. (1986), *Contemporary Theories of Knowledge*, Savage, Maryland: Rowman and Littlefield; second edition 1999 with Joseph Cruz.

[8] Cf. Pollock and Cruz (op. cit. 1999: 83) for this example and for some of the ideas in this section.

perceptual. This helps us to envisage a possible alternative justification-defeating aetiology. It is not clear that the same goes for inferential practices, considered as dispositions to move from beliefs to beliefs. Clearly there are good and bad such dispositions, but their character as good or bad is not a function of their aetiology. A good disposition might have become internalised in a "bad" way, thanks to a hectoring and dogmatic teacher, for example, and a bad disposition can be acquired on the basis of the highest authority. This does not preclude the use of empirical material in casting suspicion on, or contributing justification to, an inferential practice. We have seen such a role in the case of option negation, and there are many others. It remains the case that our inferential practices, unlike beliefs about the whereabouts of £50 notes, do not *demand* this.

Does not the admission that there are bad inferential practices mean that the attack on innocence succeeds? This would require that one would be unjustified in exercising a bad disposition, even if one has no ground to suspect it. This opinion is not supported by the illustrations, nor, as far as I am aware, by anything else. If the disposition is "bad" then it will not transmit justification, or will not result in a correct judgement of consistency or inconsistency; but that is not the same as saying that the agent was not justified in exercising the disposition. Just as one can be justified in believing a falsehood, treating as true what is in fact false, one can in principle be justified in behaving as follows: exercising a disposition which does not in fact transmit justification as though it were one which did, or exercising a disposition which does not in fact lead to a correct verdict about the consistency status of some propositions as though it did.

Turning to the kind of case envisaged in (**Error! Reference source not found.**), the present discussion relates to primitive inferential practices, ones which are not justified in terms of others. This means that no parallel example can be forthcoming for the inferential practices at issue here.

Perception provides clear examples of "cognitive mechanisms", like the visual and auditory systems, so we can make good sense of their failing to function properly. Moreover, various kinds of failure can be manifest. Cases of manifest failure are no threat to innocence, for the manifest failure constitutes a ground for suspicion. In the case of non-manifest failures, we can be drawn by one of two analogies: since non-manifest failure is failure, it is justification-defeating; alternatively, since it is not manifest, it leaves justification intact. Defenders of unqualified innocence will take the second analogy, and opponents the first. But in the case of inferential practices, the scene is different. We have no, or very little idea, about the cognitive processes which underlie inferential practices. There are no analogues of sense-organs or appropriate operating conditions or appropriate subject matters (for logic is topic neutral). There are no cases in which failure of a mechanism underlying good reasoning is manifest independently of bad upshots: nothing corresponding to a numb bodily part or, for vision, the absence of light. To suppose that the mechanism responsible for some inferential practice is malfunctioning is no different from supposing that the practice is a bad one. In this case, we can judge only by ends and not by means. But I have already said that I see no reason for thinking that to engage in a bad practice is eo ipso to reason in an unjustified way, any more than any party to this discussion would think that a false belief is eo ipso unjustified.

I suggest, therefore, that innocence needs no qualification in its application to logic, that is, to our inferential practices: whatever one may think about beliefs of the kind which should stem from memory or perception, aetiology does not figure as a condition for the justification of our primitive inferential practices. For these, innocence is a correct principle: we are justified in engaging in the practices we in fact engage in, provided that we have no ground to suspect them. In considering alternative logics, we do typically provide grounds for suspecting some

inferences in which we engage: if those suspicions are not laid to rest, our justification is undermined. In providing "grounds" for suspicion of some inferential practice, we thereby engage in others. Following this route, we cannot cast suspicion on all of them at once, so we are not engaged in a self-defeating enterprise.

R. M. Sainsbury
King's College London
mark.sainsbury@kcl.ac.uk

Further Reading

On logic and reasoning:
Harman, Gilbert (1986). *Change in View*. Cambridge, Mass: MIT Press.

On free logic:
Burge, Tyler (1974). "Truth and singular terms." *Noûs* 8, 309–325. Reprinted in Lambert, Karel *Philosophical Applications of Free Logic,* Oxford, New York, 1991, pp. 189–204.

On conditionalizing disjunctions:
Edgington, Dorothy (1995). "On conditionals." *Mind* 104.414, 235–329.

On non-classical logics:
Haack, Susan (1974). *Deviant Logic: Some Philosophical Issues*. Cambridge: Cambridge University Press.

On justification
Dummett, Michael (1978). "The justification of deduction". In his *Truth and Other Enigmas*. London: Duckworth.
Field, Hartry (1996). "The aprioricity of logic." *Proceedings of the Aristotelian Society* 96, 359–379.
Pollock, John L. and Joseph Cruz (1999). *Contemporary Theories of Knowledge*. Lanham: Rowman and Littlefield. Second edition; first edition with Pollock as sole author 1985.