Conditionalization and Belief *De Se*

Darren BRADLEY[†]

ABSTRACT

Colin Howson (1995) offers a counter-example to the rule of conditionalization. I will argue that the counter-example doesn't hit its target. The problem is that Howson mis-describes the total evidence the agent has. In particular, Howson overlooks how the restriction that the agent learn 'E and nothing else' interacts with the *de se* evidence 'I have learnt E'.

Conditionalization

Bayesian confirmation theory says that agents should update by conditionalization when a new piece of evidence is learnt:

> *Conditionalization* Suppose an agent has prior probabilities $P(H_i)$ at t0. If the agent learns some particular proposition, E, and nothing else between t0 and t1, then her t1 probabilities should be $P(H_i|E)$, where $P(E) > 0$.

Note the restriction that the agent learns 'E and nothing else'. It might seem hard to ever satisfy this requirement, for whenever E is learnt, the agent is also in a position to learn 'I learn E'.[1] Nevertheless, 'E' and 'I learn E' are distinct propositions, and conditionalization only applies when E expresses the total evidence the agent has learnt. I will argue that a purported counter-example to conditionalization, due to Colin Howson, overlooks this point.

Howson's argument

I will quote Howson's argument in full. He claims that you can be consistent

> *even if you now know what $P_2(A)$ will be in the event of E's being true, and know that it differs from $P_1(A|E)$*, as the following example, due to Richard Thomason shows. A husband announces 'if my wife is unfaithful, I shall never know'; – the wife being known to be an expert in deception. The corresponding conditional probability he ascribes to his not knowing that his wife is unfaithful [A], given his wife's infidelity [E], is presumably 1 or near 1. Yet learning that his wife was unfaithful he could

---

[†] Philosophy Department, City College of New York, NY, USA; Email: dbradley@ccny.cuny.edu

[1] These higher-order distinctions are paid more attention by epistemologists than by formal epistemologists (see Alston 1980). Stalnaker (1984) notes the distinction, and it is incorporated to varying degrees by Bar-Hillel and Falk (1982), Freund (1965), Horwich (1982), Korb (1994), Nickerson (1996), Shafer (1985) and especially Bostrom (2002). But see Hutchison (1999) for the only detailed and systematic discussion with regard to conditional probability. Also see fn. 3.

scarcely consistently assign probability close to 1 to not knowing what he has just learnt, contrary to the precedent laid down in [the principle of conditionalization] (Howson 1995, 9, italics original).

This has become known as a Thomason case. Letting P(A|E) express the conditional probability of A given E, and letting $P_E(A)$ express the probability of A after E is learnt, Howson claims that:

$$P(A|E) = \text{High}$$

and

$$P_E(A) = \text{Low.}$$

So this man violates conditionalization, yet he is not irrational in any way. Howson concludes that conditionalization is not a requirement of rationality. I will argue that the (rational) man would in fact obey conditionalization if his evidence were correctly described. The reason is that the story is ambiguous between whether he learns E and nothing else, or whether he learns 'E and I learn E'.

Higher-order de se *beliefs*

A belief *de se* is a belief about oneself. A higher-order belief *de se* is a belief about one's beliefs. When a husband learns that his wife is unfaithful, he generally learns a higher-order belief *de se* as well – he learns that he learns his wife is unfaithful. But he doesn't have to. Suppose he has low powers of introspection and literally learns nothing but E. If so, $P_E(A) = P(A|E)$, so conditionalization will not be violated. Both will have a very high value, reflecting the husband's earlier statement that if his wife were unfaithful he wouldn't know. Even though the husband has learnt that his wife is unfaithful, he will assign a high probability to (A) 'I don't know my wife is unfaithful'. There is nothing inconsistent about this. 'My wife is unfaithful' is a statement about his wife. 'I don't know my wife is unfaithful' is a statement about his knowledge. There is no necessary connection between the two.

   Perhaps having a high level of coherence between first-level and second-level beliefs is a constraint on rationality. Let's grant that it is. Then if the rational man learns that his wife is unfaithful, he also learns that he learns his wife is unfaithful. But then the husband doesn't learn 'E and nothing else' – he learns 'E and I learn that E'. So

$$P_E(A) = P(A|E)$$

but

$$P_{\text{E and I learn that E}}(A) \neq P(A|E).$$

We still have

$$P_{E \text{ and I learn that } E}(A) = P(A|E \text{ and I learn that } E),$$

which will be virtually 0. So there is again no violation of conditionalization. Howson's counter-example fails because it equivocates on whether the evidence is 'My wife is unfaithful' or 'I have learnt my wife is unfaithful'.

There is a general issue here about how the Bayesian should respond when we take into account *de se* evidence regarding what the agent himself has learnt. It is widely agreed that the agent should conditionalize on his total evidence (Good 1967), which can be understood as the strongest proposition of which the agent is certain. But *de se* evidence appears to generate a problem for the concept of total evidence. Hájek (2007) puts the point as follows[2]:

> When we go beyond toy examples it is unclear whether there really is such a thing as 'the strongest proposition' of which one is certain. [For any proposition X that purports to express the total evidence] . . . it seems that you also learned a potentially infinite set of *de se* propositions: 'I learned that X', 'I learned that I learned that X' and so on. Perhaps, then, your total evidence is the infinite intersection of all these propositions, although this is still not obvious – and it is certainly not something that can be represented by a sentence. . . . which is finite in length (Hájek 2007, 19).

Is this a problem for Bayesians? If we are modelling ideal agents (assuming a very strong understanding of 'ideal'), then there is no problem with them completing an infinite number of tasks and updating on an infinite number of propositions. If the agents are non-ideal however, then eventually there will be iterations of 'I learned that I learned . . . X' that will not be included in the total evidence conditionalized upon. But the agent will only come to the wrong conclusion under very specific conditions; there must be an iteration of 'I learned that I learned . . . X' that is not conditionalized upon yet is probabilistically relevant to some hypothesis of interest. And this will only happen in the most contrived cases. For example, 'I learned that I learned that X' will be positively relevant to the hypothesis that I am highly reflective. For most hypotheses however, your own higher-order beliefs will be irrelevant. Conditionalization applies to eternal propositions, which, if true, are always true. But 'I learn X' changes from being false to being true. For a more familiar example, 'It is Monday' changes from being false to being true. As a result, agents can come to believe 'It is Monday' by tracking the truth-value, as opposed to discovering new evidence. This is indeed a type of belief change that cannot be

---

[2] This issue is also raised by Chihara (1994). Interestingly, Howson (1997) denies that this creates a problem for Bayesians, saying that although there is no reason in principle why the semantic ascent of 'I learn that I learn that . . .' couldn't continue forever, 'in practice there seems no good reason to deny that sometimes we simply learn things without simultaneously learning that we have learned them. Of course, if questioned, one may reflect that one knows what one knows, and add this as a new layer of knowledge, and the propositions describing it can be added to the domain of one's probability function'. This is compatible with my response that the higher-level knowledge will only be relevant in the most contrived cases.

modelled by conditionalization. For discussions of this problem see Bradley (forthcoming), Meacham (2008), Titelbaum (2008) and Schwarz (ms).

Howson's argument against conditionalization is different. It is based on a Thomason case, where the fact that E is learnt is relevant to the hypothesis. But the principle of conditionalization includes the condition that 'E and nothing else' is learnt. The purported counter-example trades on smuggling in higher-order *de se* beliefs about the agent's own evidence, whereby the condition fails to be satisfied. So the purported counter-example to conditionalization fails.

References

ALSTON, W.P. 1980, 'Level Confusions in Epistemology', *Midwest Studies in Philosophy* **5**, pp. 143–145.

BAR-HILLEL, Y. and FALK, R. 1982, 'Some Teasers Concerning Conditional Probability', *Cognition* **11**, pp. 109–122.

BOSTROM, N. 2002, *Anthropic Bias: Observation Selection Effects in Science and Philosophy*, New York: Routledge.

BRADLEY, D. forthcoming, 'Self-location is no problem for conditionalization', Synthese.

CHIHARA, C. 1994, 'The Howson–Urbach Proofs of Bayesian Principles', in: E. Eells and B. Skyrms, eds, *Probability and Conditionals*, Cambridge: Cambridge University Press, pp. 161–179.

FREUND, J. 1965, 'Puzzle or Paradox?' *American Statistician* **19**, pp. 29–44.

GOOD, I.J. 1967, 'On the Principle of Total Evidence', *The British Journal for the Philosophy of Science* **17**, pp. 319–321.

HÁJEK, A. 2007, 'The Reference Class Problem is Your Problem Too', *Synthese* **156**, pp. 185–215.

HORWICH, P. 1982, *Probability and Evidence*, Cambridge: Cambridge University Press.

HOWSON, C. 1995, 'Theories of Probability', *British Journal for the Philosophy of Science* **46**, pp. 1–32.

HOWSON, C. 1997, 'On Chihara's The Howson–Urbach Proofs of Bayesian Principles', *British Journal for the Philosophy of Science* **48**, pp. 83–90.

HUTCHISON, K. 1999, 'What are Conditional Probabilities Conditional Upon?', *British Journal for the Philosophy of Science* **50**, pp. 665–695.

KORB, K. 1994, 'Infinitely Many Resolutions of Hempel's Paradox', in: R. Fagin, ed., *Theoretical Aspects of reasoning and Knowledge*, San Francisco: M. Kauffman, pp. 138–149.

MEACHAM, C. 2008, 'Sleeping Beauty and the Dynamics of De Se Beliefs', *Philosophical Studies* **138**, pp. 245–269.

NICKERSON, R. 1996, 'Ambiguities and Unstated Assumptions in Probabilistic Reasoning', *Psychological Bulletin* **120**, pp. 410–433.

SCHWARZ, W. (ms), '*Changing Minds in a Changing World*'.

SHAFER, G. 1985, 'Conditional Probability', *International statistical Review* **53**, pp. 261–277.

STALNAKER, R. 1984, *Inquiry*, Cambridge, MA: Bradford Books.

TITELBAUM, M. 2008, 'The Relevance of Self-Locating Beliefs', *Philosophical Review* **117**, pp. 555–605.