Evaluating Background Independence

Robert G. Hudson
Department of Philosophy
Algoma University College
Sault Ste. Marie, Ontario

hudsonrg@hotmail.com

Abstract

Over the years, a number of philosophers have advocated the following epistemic principle on observations (which I call 'background independence'): when using observational data to test some theory, use data that do not, in themselves, presuppose the truth of the theory under test. Using as the focus of my criticism Sober's recent presentation of this principle (Sober 1999), I argue that background independence is not the epistemically valuable principle some philosophers think it is. This leads me to consider why philosophers have been attracted to this principle in the first place, and my suspicion is that they have equivocated between two versions of the principle, one which is epistemically reasonable (but which is not a constraint of observations after all, but on the observation−theory evidential link), and the other which is not epistemically reasonable (but sometimes confused with the former version of the principle, giving it the appearance of being reasonable). Subsequently I argue that, by reflecting on the Duhem−Quine nature of experimental testing, one arrives at deeper understanding of what is at stake with background independence. I conclude by countering an objection to my view, an objection based on the occasional need for blind tests.

Evaluating Background Independence

1. Introduction. A pressing philosophical question is the following: What guarantee do we have for the reliability of observational (empirical, perceptual) data? The theory−ladenness of observation has cast a shadow on this question; absent any direct perceptual access to the world, it is highly possible that what we perceive is preformed by our theoretical beliefs about the world. This is problematic for we hope to use observation as a foil for theoretical preconceptions. This will not occur in many important cases if observations themselves have a theoretical bias.

My task here is to evaluate and criticize one popular attempt to resolve this philosophic problem. This attempt I call 'background independence'. Here is how background independence works. Suppose we wish to empirically test some theory. What observations can we use to reliably test this theory? More precisely, which observations can we use to potentially falsify this theory? Background independence suggests the following strategy: given that observations will be theory−informed, come what may, use those observations that do not, themselves, presuppose the truth of the theory under test.

On the face of it, this is a reasonable strategy, and it has been upheld as such by many philosophers: Greenwood (1990), Kosso (1988) and (1989), Hacking (1983), Carrier (1989), Wylie (1990), plus others cited in Hudson (1994). Most recently, it has been lauded in Eliot Sober's recent 1999 Presidential Address to the Central Division of the American Philosophical Association (Sober 1999). To give the flavor of Sober's

view, I quote from that address. Sober comments:

> observations are able to help answer a question concerning which of several competing hypotheses is most plausible only to the extent that we are able to decide which observation statements are true without first having to know which of the competing hypotheses is true. It is in this sense that observations must be 'theory−neutral' – they should be neutral, relative to the competing theories under test; they need not be neutral in any absolute sense. . . . Other theories can be used to justify various observation statements, but the theories under test cannot be presupposed by the observation statements that are used to test those very theories. (p. 52)

One wonders: what could be wrong with so intuitive and obvious a constraint on observations? My goal here is to raise some questions about this constraint; by considering these questions, I believe we can come to a deeper understanding of the evidential significance of observations.

2. Criticizing Background Independence. To start our discussion of background independence, let me give a simplistic example of how an observation, presumably by being laden with the theory under test, cannot succeed in objectively testing this theory.

Let us imagine an individual who has a drastically paranoiac view of the world. We might describe such a person as having a theory about the world, the 'paranoid's theory', expressible as the following claim: "Everyone is trying to kill me". Indeed, let us imagine that this individual is so paranoid that any observation he makes is informed

by this theory. For example, at dinner he observes his wife shaking 'salt−flavored poison' onto his food (not just plain salt!), in the morning he observes the mailman 'hooking up an explosive device' on his mailbox (not just delivering a phone bill!), and throughout his day he makes similar observations. Still, our paranoid is rational enough to ask himself the question: "What observational evidence do I have for my paranoid view of the world?" "Well," he thinks, "I saw my wife trying to poison me, the mailman trying to blow me up, ..." Here is a classic case of background dependence, as Sober and the others seem to construe it. The paranoid is interpreting his observations with precisely the theory under test, thereby precluding the possibility that the observational evidence will refute his theory.

But now let us diagnose what has gone wrong with the paranoid's empirical strategy. Admittedly, his observations do not confirm his paranoid theoretical perspective. But neither do his observations confirm nor observationally test any other theoretical perspectives, even those not presupposed by the paranoid. For, essentially, what is going wrong with the paranoid's observations is that they are failing to track the nature of the world – he is observing poison and explosives when they're not there – so for even those theoretical perspectives not presupposed by the paranoid these observations lack evidential significance. To demonstrate that it is not assuming what one is trying to prove that is the problem, but rather the unreliable nature of the observations, consider what our appraisal of the situation would be if, miraculously, the paranoid were dead right about his observations, and consistently so – that his wife *was* trying to poison him, the mailman *was* trying to blow him up, and so on. These

observations would surely succeed in confirming his paranoid perspective, despite the fact that this perspective is their source.

It is curious, then, why so many philosophers have been drawn to background independence as an epistemic constraint on observations. To get a handle on what has motivated them, consider some further comments from Sober (1999). He says,

> we may be inclined to treat the mass of the earth as something that we observe when we use the measured mass to answer a question about something else. But if we are trying to figure out what the mass of the earth is, we may want to reserve the term "observation" for a different class of statements, ones whose truth values we can ascertain without already knowing what the mass of the earth is, and which allow us to discriminate among different estimates of that quantity. (p. 52)

What is Sober suggesting here? It is not entirely clear – he is talking about 'ascertaining the truth−value of observations (of the mass of the earth)' and, to me, such an activity does not sound like the process of observation (nor even like the process of uttering or writing observation statements). 'Ascertaining a truth value' sounds more like an activity one engages in *after* observations or statements are recorded. Thus it seems that when Sober is describing the process of observation, he is describing an evaluative process, one in which a scientist is deciding how to interpret the observations. In such an evaluation, all can agree that if the observations being considered are posed as testing a theory, it would be unwise and counterproductive to assume *this* theory in the evaluation. To illustrate this point, let me consider a very simple example. Suppose, by utilizing a

theory T, a scientist predicts that observation O will occur, and that an experiment is set up to test whether O occurs. After running the experiment, the scientist wishes to decide whether the produced result confirms or disconfirms the theory, and for this purpose uses T in evaluating the significance of the generated result. And let us suppose that, to this end, she adopts the following logic: if O occurs she concludes that the experimental result confirms T (on the grounds that T predicted O); on the other hand, if O does not occur, she dismisses the accuracy of the observed result (on the grounds that the observed result runs contrary to what T predicts). Clearly, this is a regrettable, circular testing procedure – and if prohibiting such a possibility is all background independence amounts to, I would not hesitate to affirm it, either.

So, phrased as, "in testing a theory using an observation, do not use *that* theory in evaluating the evidential significance of this observation", background independence seems unexceptionable. And perhaps it is this epistemic constraint that Sober and the others have in mind in their discussions. Yet, note the following: the proposed constraint is not a constraint on *observations*, but only a constraint on theory−testing, on the *observation−theory evaluative link*. In other words, there is nothing in background independence, as here construed, that prohibits using the theory under test in retrieving observations, such as in using the theory in the design of experimental apparatus, in recommending experimental strategies or in suggesting why a particular environmental context will be useful in generating informative observations. Should background independence be extended to cover these further cases? One way of interpreting my above discussion of the paranoid was to suggest that applying background independence

to these further cases is unnecessary, even irrelevant, so long as the theoretical resources we use to generate observational evidence assure the reliability of this evidence.

How then do those philosophers who assert the value of background independence interpret this principle? Do they adopt what I take what I take to be the incorrect reading of this principle, that is, one which proscribes assuming the theory under test in suggesting experimental strategies, designing apparatus and so on? One finds their writings on this issue to be a bit vague. Consider again Sober:

> Other theories can be used to justify various observation statements, but the theories under test cannot be presupposed by the observation statements that are used to test those very theories. (p. 52)

I ask: what does it mean here for an observation to 'presuppose a theory' (leaving aside the question of whether it is the theory under test)? A similar locution occurs in Greenwood (1990):

> although it is true that our currently accepted ontological beliefs are (or ought to be) determined by our best theories, the observations which confirm our best theories *do not generally presuppose them*. (p. 560; Greenwood's italics)

Again: what does it mean here for observations to 'presuppose a theory'? Another background independence advocate, Peter Kosso, suggests that an observation is theory−dependent "in the sense that there is some collection of theories which account for the information one claims to receive in the observation, information that functions as evidence in science" (1989, p. 250). And, for him, epistemic objectivity in testing is ensured if the collection of theories "which account for the information one claims to

receive in the observation" is logically independent of the theory under test. We ask: what does it mean for a theory to 'account for the information one receives in the observation'?

After examining their writings, I do not believe Sober, Greenwood and Kosso are clear enough on what it means for an observation to assume the theory under test, so I leave aside a closer examination of their views. Instead, I propose examining the work of Martin Carrier who offers some illumination on this issue, following on the efforts of Joseph Sneed. Carrier discusses in his (1989) the epistemic problems raised by so−called Sneed−theoretical terms, terms for which "*all* means of determining the truth−value of statements involving the [term] presuppose the truth of the laws of the theory in question (p. 411; quoted from Sneed 1979, p. XVIII; Sneed's italics). Carrier's point is that some observation statements, because they are Sneed−theoretical and so presuppose the theory under test, cannot function in testing such a theory. Accordingly, he recommends that we avoid observational situations involving Sneed−theoretical terms which is nothing less than a recommendation to adopt the principle of background independence. But again, to give value to this principle, one needs to be informed about what it means to 'presuppose the truth of the laws of the theory in question', and here Carrier/Sneed provide the following account. They consider the case of classical particle physics and, in particular, observations of masses and forces used in testing Newton's second law. These observations are problematical, on Sneed's view, since "all procedures for measuring masses and forces should make use of this second law" (Carrier 1989, p. 412). In other words, observations of masses and forces 'presuppose' Newton's second law

theory in that procedures generating these observations involve calculations that use as premises statements of this law. In his 1989 paper, Carrier's view is that if these observations 'presuppose' Newton's second law in this sense, it becomes impossible to test Newton's second law using mass and force observations. Accordingly, he explores ways to measure masses and forces without relying on the second law as a calculational device (1989, p. 417).

This takes us a little farther in understanding what it means for an observation to presuppose a theory – a theory is presupposed in that certain of its laws are used as premises in calculations moving from received data to observations. Let us now explore whether Carrier's concern over Sneed−theoreticity is justified. I think not as the following abstract example illustrates. Suppose we observe that an object has mass m, where this observation involves a procedure which utilizes Newton's second law in the context of a calculation. And let us imagine that, given certain other theoretical commitments we possess, ones *different* from Newton's second law, we expect an object of mass m to have certain observable effects. Finally, suppose these observable effects do not come about. What should our conclusion be here? We have at least two alternatives – first, we can challenge these other theoretical commitments; or if these commitments are dear to us, we can question whether our calculation of mass was performed correctly. If the latter, why could we not be prompted to blame the second law for this error? If we did, it would follow that our mass measurements, produced through the assumption of the second law, have in fact induced us to question the second law. That is, background dependence has not, in any way, shielded the theory informing

observations from criticism. Thus, it is not clear to me what motivates Carrier's concern over Sneed−theoretical terms, why, in effect, he advocates background independence.

Now, abstract though the above example is, surely the sort of situation it describes is quite possible and that in this way theory−induced observations can be used to question the inducing theory. Where there are untoward experimental results, the issue always arises concerning where to pin the blame. In particular, one always has the option to question those theories underlying the observational or experimental process. In other words (and here Greenwood (1990) is on the right track), the circularity problem addressed by background independence is resolved simply by the Duhem−Quine nature of the experimental justification of theories, by the vulnerability to rejection of any theory, whether the one under test or those used as auxiliaries. For despite the fact that one's observations are informed by a theory T, so long as an experimentalist is willing to question T, given untoward results, there is no need to be concerned about vicious circularity in one's testing procedure. The crux is the experimenter's *attitude* toward the testing situation, to wit, is she genuinely open to questioning the theory informing the observations? If not, then this is the sort of circularity we should be concerned about.

Accordingly, it would be inappropriate for us, as philosophers, to recommend to scientists that they not assume in the observational process the theories that are under tests. Instead, we should be telling them probably what they already know, that their goal is to produce reliable observations, even if these assume the truth of the theory being tested.

I plan in the next section to consider a rebuttal to my attack on background independence. Before that work, though, let me briefly connect my approach to two different approaches that philosophers have taken in the literature in responding to the threat of observational circularity.

To begin, Brown (1994) makes the following suggestion: where observations are generated using calculations applying theory T, these observations could nevertheless serve to test and even disconfirm T, thus averting the spectre of circularity, if they are combined with *other* observations and further alternative calculations are made with these other observations that, as it turns out, show T to be false. He illustrates this strategy using as a case example the calculation of the velocity of celestial objects using earth−bound telescopes (see Brown 1994, pp. 408−409). One can use these telescopes to measure the change in angle between two distant celestial objects; and along with various theoretical assumptions, including relativity theory, one can calculate the recession velocities for these objects on the basis of this change in angle. Such recession velocities, as calculated, cannot attain a value greater than the speed of light since they were arrived at using relativity theory as an assumption. (Here is the problem of observational circularity.) However, further calculations using these recession velocities, along with the Hubble constant, the measured change in angle between the celestial objects plus – and here are the *other* observations – the known time between the two measurements used to determine this change in angle, could yield velocities for light that violate relativity theory. In this way Brown believes we can avert observational circularity.

Now, I do not doubt that this approach works. Nevertheless, I do not believe these other observations need to be made. If the initial observations, it turns out, run counter to certain other theoretical preconceptions we possess (which of course differ from those preconceptions informing our observations), we have the option of questioning the theory informing our observations, and even overturning it on the basis of these observations. Again, it is the Duhem−Quine nature of testing, combined with an express willingness on the part of an experimenter to question those theories informing the observations, that resolves the circularity problem. It does not really matter whether the theory informing the observation is questioned because of the conflict of this observation with other theories (my approach) or the conflict of the informing theory with other observations (Brown's approach). One has problematic circularity, in any case, whenever a theoretical perspective is upheld, come what may.

A different approach to these issues is adopted by Franklin et al (1989) and by Chalmers (1999). The sort of case Franklin et al (1989) consider is this: suppose we are testing the hypothesis, "objects expand as their temperature increases", with a mercury thermometer (see p. 230). There is a problem here, they note, in that "such a thermometer depends on that same hypothesis for its proper operation" (ibid.). Chalmers considers a similar example involving the measurement of the deflection of a current−carrying coil suspended between the poles of a horseshoe magnet, where the measurement of the current utilizes an ammeter that contains, too, a current−carrying coil suspended between the poles of a magnet (see Chalmers 1999, pp. 38−39). He takes this

to be a case where a theory is "appealed to in order to judge the adequacy of experimental results, and those same experimental results are taken as evidence for the [theory]" (p. 38). What do Franklin et al and Chalmers recommend as a way to bypass the troublesome circularity in these cases? Franklin et al suggest calibrating the mercury thermometer with a constant volume gas thermometer, thereby "[transferring] the theory of the apparatus from one theory to another, thus separating the theory of the apparatus from the theory under test" (p. 230). And although Chalmers does not recommend a specific alternative approach, he emphasizes that what is needed is a "method of measuring the current in the circuit that [does] not employ the deflection of a coil in a magnetic field" (p. 39). For both Franklin et al and Chalmers, then, their recommendation is that when investigating particular empirical phenomena (e.g., the expansion of physical objects with an increase in temperature or the deflection of current−carrying coils suspended between the poles of a magnet), one should not employ an experimental apparatus that exploits the same phenomena.

Now, they might be right about this. But it is irrelevant to the problem we are addressing in this paper. Our problem is this: can one legitimately use a theory in the observational process (such as in performing certain calculations) if this theory is the one being tested? The problem Franklin et al and Chalmers are considering is this: can one legitimately use an empirical phenomenon in the observational process (e.g. the expansion of physical objects with an increase in temperature) if this phenomenon is the one being investigated? These are separate problems. The problem we are considering in this paper is called 'background dependence'. And although I do not have a name for

the problem Franklin et al and Chalmers are considering, it is a different problem from ours, and as indicated by the passages quoted in the previous paragraph, they are confusedly equating it with background dependence.

3. Blind Tests. I have been arguing that, in testing theories using observations, one need not worry about informing those observations with the theory under test. Let me now consider an objection to my view.

My dismissal of the epistemic significance of background independence (as a constraint on observations) can be attacked in the following way. It is common knowledge that, occasionally, experimentation in social research and clinical medicine requires the use of blind trials. Where a researcher is experimenting with human subjects, a single−blind trial conceals important facts from the subjects participating in the study. A double−blind trial goes further and conceals important facts from *both* these subjects and the experimenter. So, for example, in testing the effectiveness of a drug through a comparison of the effects of this drug on those receiving it with those receiving only placebos, an experimenter in a double−blind trial will be unaware of who receives the genuine drug and who does not. Why is it prudent to conceal from an experimenter information about the subjects of the study? The reason is that an over−zealous researcher, perhaps maintaining a theory that predicts improvements in patient health and keen to see such improvements, will 'see' these improvements even if they are not present One wonders then: does not this sort of phenomenon, dealt with by recourse to double−blind tests, show that background dependence is indeed a hazard, that we need to

be wary of people importing their theories into observations, especially where the theories imported are the ones under test? To that extent, background independence as an epistemic strategy (perhaps as an alternative to blind tests) must have some merit, despite what I have said above.

In response: there is no doubt that in some situations people will see what they want to see in the data. These are situations where the determination of an observational result is highly interpretive. Social researcher Earl Babbie puts the point in the following way. As he comments,

> in social science experiments, or in medical experiments, the danger of experimenter bias [and so the need for blind tests] is . . . reduced to the extent that the operational definitions of the dependent variables are clear and precise. (Babbie 1998, pp. 236−237)

Or again, medical statistician Theodore Colton makes these remarks:

> the need for a double−blind trial depends on the nature of the response used to assess the results. The more subjective the outcome may be, the greater is the rationale for a double−blind trial. (Colton 1974, p. 264)

Now such highly interpretive experimental situations are clearly problematic – the recorded observations will quite often be the product of the experimenter's preconceptions and, as a result, there is a proper concern for the reliability of such observations. However, this sort of situation is problematic *whether or not* the experimenter's preconceptions are those that are under test. That is, the problem here is not that one is, specifically, interpreting observations using the theory under test, but

that one has too much leeway in interpreting observations to begin with. Accordingly, adopting background independence does not solve here the problem since the unfortunate fact is that whatever dubious preconceptions the experimenter harbors, whether under test or not, these preconceptions effectively inform the observations and quite possibly damage their reliability. Let me conclude by putting this point in opposite fashion: there is no need to be concerned that our theories of choice are informing our observations, even if these theories are the ones under test, so long as we can count on the resultant observations to reliably depict the nature of the world.

References

Babbie, E. (1998), *The Practice of Social Research*, 8th edition. Belmont, California: Wadsworth Publishing Company.

Brown, H. I. (1994), "Circular Justifications", *PSA 1994*, vol. 1, pp. 406−414.

Carrier, M. (1989), "Circles Without Circularity", in J. R. Brown and J. Mittelstrass (eds.), *An Intimate Relation*, Dordrecht: Reidel, pp. 405−428.

Chalmers, A. (1999), *What is this thing called Science?*. Indianapolis: Hackett Publishing Company.

Colton, T. (1974), *Statistics in Medicine*. Boston: Little, Brown and Company.

Franklin, A. et al (1989), "Can a Theory−Laden Observation Test the Theory?", *British Journal for the Philosophy of Science* 40: 229−231.

Greenwood, J. (1990), "Two Dogmas of Neo−Empiricism: The 'Theory−Informity' of Observations and the Quine−Duhem Thesis", *Philosophy of Science* 57: 553−574.

Hacking (1983), *Representing and Intervening*. Cambridge: Cambridge University Press.

Hudson, R. (1994), "Background Independence and the Causation of Observations", *Studies in History and Philosophy of Science* 25: 595−612.

Kosso, P. (1988), "Dimensions of Observability", *British Journal for the Philosophy of Science* 39: 449−467.

Kosso, P. (1989), "Science and Objectivity", *Journal of Philosophy* 86: 245−257.

Kuhn, T. S. (1996), *The Structure of Scientific Revolutions*, 3rd edition. Chicago: University of Chicago Press.

Sneed, J. (1979), *The Logical Structure of Mathematical Physics*, 2nd edition. Dordrecht: Reidel.

Sober, E. (1999), "Testability", *Proceedings and Addresses of the American Philosophical Association*, 73: 47−76.

Wylie (1990), "Varieties of Archeological Evidence: Gender Politics and Science", presented at the Eastern Division meeting of the American Philosophical Association, December 1990, Boston, Mass.