Scientific Anti-Realism and the Epistemic Community

William Seager

University of Toronto

The ability to observe is the ability to reliably detect, but that is not all observation is. A thermometer reliably detects temperature yet does not observe the temperature, whereas I do, even though in terms of reliability I cannot match the thermometer. An observation is detection accompanied by active classification and, typically, the subsequent formation of opinion. Even when we say of an animal that it can see something we mean more than that it reliably detects things of a certain sort but also that it deals with such things *as* objects, gauging their usefulness or lack thereof in relation to its interests and to other aspects of the world. This crucial addendum to the notion of observation can be enshrined in slogan: observers are potential believers. This marks the difference between mere detecting mechanisms and observers.

The contemporary scientific anti-realism of one such as van Fraassen, which is developed in detail under the label of "constructive empiricism" in van Fraassen (1980), can be characterized in terms of detecting mechanisms. Note first that the term 'detection' is a misnomer in cases where these mechanisms are detecting what is not observable. For obviously 'x detects y' entails that y exists and it would be an absurdly facile route for the realist to say that since film can detect x-rays, the question of whether x-rays exist is already settled. Since the anti-realist quite rightly wishes to maintain that questions regarding the existence of unobservables are not so easily settled, some other tack on 'detection' will be required. The account is, in rough form, that science aims at predicting and explaining the observable phenomena, some of which will be the operation of so–called detecting mechanisms. In the example of the film, physics predicts that under certain circumstances film will become exposed in an apparently perfectly dark room. Such a prediction is entirely within the realm of the observable. Furthermore, the reason film is said to *detect* x-rays is that photographic technicians are, in van Fraassen's word, 'immersed' in a theory from the standpoint of which x-rays darken the silver halide grains in the film, which interaction is called exposure. Such immersion is unseemly however for the philosopher of science who steps back from the admittedly engrossing activity of science and who then sees that in fact the epistemic value of a theory consists entirely in finding and predicting connections among observable phenomena. Naturally, theories will have other *virtues* on the anti-realist account, such as simplicity, elegance, consilience with other theories and the like, but the point here is that these will not be *confirmatory* virtues. For more on this see van Fraassen (1983).

182

One species of detecting mechanism is the human observer, and the question naturally arises whether this aspect of the anti-realist view of science leads to a treatment of the observer as a *mere* detecting mechanism. In some ways, the answer is yes, for science will specify the limits, in general terms of size, brightness, loudness, etc, regarding what the human observer can reliably detect. But in a much more important sense, the answer is no. If observers were no more than detectors then the very distinction between observable and unobservable would, if not collapse, seriously mutate. This is because if humans were treated as mere detectors the standpoint which divides detectors into those responding to observable and those (putatively) responding to unobservable items would disappear. Why select one set of *detectors* as privileged? At best, on this view, observability becomes a two termed relation, x is observable to y, with no ground for selecting a special y to mark out the 'really' observable.

When we make, as we do, our selection of ourselves as this special y, thus restoring to 'observability' its monadic status, we do so on the basis of our standing as potential believers, not as mere detectors. We *know* the range of objects we can detect (more or less) and compared to us mere detectors can be treated as just another aspect of the empirical world rather than as carriers or propounders of a view of the world which itself deserves credence.

The point here is not that we fail to stand in an epistemic relation to detectors for obviously we do–they indisputably provide us with information; the problem is rather that they stand in no epistemic relation to themselves. They do not and cannot acquire beliefs on the basis of their detections and thus cannot possess a view of the world which we must take into account *as* a view of the world. Thus, we need only take account of their activity as more phenomenal grist for the mill of science.

We have so far divided detectors into two classes, those which are potential believers and those which are not. For convenience label the former 'smart detectors,' the latter 'dumb detectors.' The only smart detectors we know of are human beings and possibly some higher animals. There is a question with regard to animals as to whether or not they *are* smart detectors, since they do not interact with us in ways that unambiguously reveal the answer. Such ways are many: making reports to each other, doing experiments with each other, conjecturing, refuting, etc. These activities bind us into, to use the term countenanced by van Fraassen himself, an *epistemic community*.

The concept of an epistemic community deserves further attention. The epistemic community can be considered from an ideal or a realistic standpoint–as what it ought to be or what it is. From the former viewpoint the epistemic community is a *regulative ideal*–it presents us with a picture of a community of seekers after truth constrained only by the limitations imposed by their finite epistemic capabilities but without any limitations set by their relationships with one another. Such a regulative ideal is the one the scholarly/scientific community attempts to emulate, to some extent at any rate, by such procedures as blind refereeing, free exchange of information, publication of the methods employed in experimentation, etc.

The ideal epistemic community is made up of fully rational and also sincere members, that is, creatures who never attempt to deceive one another. I will also assume that the members are fully aware of one another's epistemic abilities–the sorts of beliefs that they would acquire in given evidential situations and that they know what epistemic situation they are actually in. (This last condition is particularly ideal, but I am more interested in the epistemic principles relating the members of the community to each other than those relating them to the world.) Now, several principles seem fairly apparent about such an epistemic community. First, and trivially, every member of such a community accepts every other member as such. Each member regards any other as one capable of belief and rational thought. Given our assumption about knowledge of

epistemic abilities this means that each believes that any other has some beliefs. Let me codify this principle (and others to come) in a semi-formal way. This principle becomes:

$$P1. \quad (x)(y)(\exists P)(B_x B_y P).$$

('B_x' is taken as 'x believes that ...'; 'P' is a proposition; the x and y quantifiers range over members of the epistemic community.) A more interesting principle is:

$$P2. \quad (x)(y)(P)(B_x B_y P \rightarrow B_x P).$$

P2 is correct since our believers are fully rational and aware of the epistemic abilities of the others. If x, for example, knew that y was in a situation where a false belief would appear rational, then y would know this too and refrain from forming the belief. These first two principles can be called 'internal' principles as they crudely outline the attitude of one believer towards another. P1 says that other members of the epistemic community are indeed capable of belief. P2 says that other members are accepted as doing a good job in the formation of their beliefs or that they are 'worthy' members.[1]

There are also external principles; ones governing the interaction of believers. I consider two as basic. The first is the principle of communication. If a member of an epistemic community believes that P then he or she will communicate this to others. Codify this as:

$$P3a. \quad (x)(P)(B_x P \rightarrow x \; says^2 \; 'P').$$

("'P'" is a name of "P".) Furthermore, since we are assuming sincerity amongst our believers we have:

$$P3b. \quad (x)(P)(x \; says \; 'P' \rightarrow B_x P).$$

Another sort of external principle is one of what might be called empathy. We have, in fact, already assumed it for it states that a member of an epistemic community knows what another would believe in a given epistemic situation. Our believer knows this because he knows what *he* would believe in that situation, so:

$$P4. \quad (x)(y)(P)((B_x Fy \; \& \; B_x(Fx \rightarrow B_x P)) \rightarrow B_x B_y P).$$

It is easy to prove from these principles the following theorem: $(x)(y)(P)(B_x P \equiv B_y P)$. Such a result emphasizes how idealized our notion of epistemic community is, but the flavor of the result is correct. If we abstract from irrationality and insincerity and consider ourselves as equally competent laborers in a common enterprise aimed at the acquisition of knowledge then we ought to end up all believing the same things (and refraining from belief in common where evidence is scanty or ambiguous). Or, to put the point in another way which bears on real epistemic practice, when we differ (or appear to differ) from another in belief we must put this down to error (in assessing the epistemic situation, say), irrationality (in lacking the ability to correctly assess evidence or in believing when there is not enough evidence to warrant it) or insincerity.

This assumes that it is irrational to believe without warrant–that one should be agnostic where grounds are lacking. And, more contentiously I think, that there is some correct way of linking evidence to belief. In other words, it is the assumption that there is an inductive logic acceptable to all the rational. This assumption can be made more palatable if we recognize that it can be relativized to epistemic communities. *Within* an epistemic community there ought to be agreement about the standards of evidence and the bearings of evidence upon belief. To the extent these are lacking there will be frag

mentation in the epistemic community. In the ideal case, which we are discussing, there will be perfect agreement on these issues.

There are at least two lessons to be drawn from this exercise. First, that the members of an epistemic community ought to respect each other's epistemic capabilities. Second, that the beliefs of a fellow member of one's epistemic community warrant belief (in the ideal case) or at least serious consideration (in less ideal cases) by oneself. Such beliefs represent a part of the world picture which we aspire to draw, but which, due to the efforts of others, we need not draw for ourselves.

One is in an epistemic community if one is taken to be a smart detector, among other things. That is, one's reports are taken to really be reports, not mere behavior. They are taken as, modulo the ever present real possibilities of error and, what only a smart detector may aspire to, insincerity, status reports on the world which may enter anyone's view of the world as interpreted rather than as just sounds arriving as new phenomenal data.

To be fully integrated into our epistemic community requires above all else being accepted as rational and linguistically capable. Some who fail these criteria are accepted as shadow members, but only on the basis of their biological kinship with us. Animals are less than full members because of their inability to communicate with us, and their relatively impoverished set of concerns. Their slim claims to membership depend on our ability to interpret their behavior as rational, in the light of what we take their concerns to be.

But we can imagine beings who, with animals, possess senses which are more acute than ours, or which are sensitive to a range of phenomena we have no direct access to, and who also would be fully accepted as members of our epistemic community. That is, we would accept them as fellow explainers and not merely as another phenomenon to be explained. Such imaginings are not foreign to the philosophy of science, for a seemingly obvious argument against the anti-realist position is to imagine humans, or human-like creatures, with extended senses. For a classic example see Grover Maxwell (1962). What they observe would then be accepted as real, i.e. believed to exist. But what do such imaginings license us to believe? Only a conditional, of the form given by van Fraassen:

If the epistemic community changes in fashion Y, then my beliefs about the world will change in manner Z. (van Fraassen 1980, p. 15).

Our examination of the notion of an epistemic community should strengthen our confidence in such conditionals, and explain why they are true. To admit new members to our epistemic community is to admit new competent believers. As such, we have in effect agreed, all things being equal, to believe what they believe.

There are at least two arguments showing that this conditional is not particularly relevant to the realism debate. First, a very simple one: why should believing that if P then Q lead me to believe Q when I don't believe P? A more complicated argument is given by van Fraassen himself, which I quote:

To see [that the anti-realist is forced to accept such conditionals] as an objection to anti-realism is to voice the requirement that our epistemic policies should give the same results independent of our beliefs about the range of evidence accessible to us. That requirement seems to me in no way rationally compelling ...(van Fraassen 1980, p. 18).

Let us consider these arguments in turn.

No doubt the simple argument is correct so far as it goes, but the case is more complicated than it suggests. First, we can add the premise: the epistemic community *could* change in fashion Y. This premise will be supported by science itself. Consider as a particular example work in artificial intelligence. Science guarantees that AI is in principle possible by way of Church's Thesis via either the reduction of psychological descriptions to principles, algorithms, etc. that involve in themselves no intelligence (a strategy made famous by Daniel Dennett–see Dennett (1978)) or the direct emulation of brain processes within a machine. Since we can build a universal Turing machine such a machine is possible, under the sole caveat that the brain, at least, is amenable to mathematical description. If the brain operates by 'magic' then AI is doomed, but I take it that we have good evidence that the brain is a purely physical device.

So our present scientific understanding of intelligence, though not very extensive, allows us to infer that there could be intelligent machines with which we *could* communicate and for which we could provide senses corresponding to any detection mechanism we wish. Thus the second premise above. Now, it permits us to draw only the following conclusion: we could change our beliefs about the world in fashion Z (assuming, as seems reasonable, that such conditionals as we are considering are in some way necessary truths). Presumably, the anti-realist has no quarrel with this conclusion either.

But let us suppose that AI succeeds. Intelligent machines become common possessing a wide range of bizarre sensory abilities (e.g. some of these are equipped with 'infrared camera eyes' for seeing at night, others with 'x-ray camera eyes' for inspecting stressed aircraft parts, and so on). These creatures naturally enter our epistemic community, for by hypothesis we can enter into full and rich communication with them. They are without question smart detectors. The upshot of the story is that any anti-realists among us are now forced to admit that x-rays are as real as chairs and tables, for they are just as observable. If they do not admit this, they are not being scientific anti-realists but instead taking up some kind of *philosophically* skeptical position. I am not interested in such positions here; after all, the anti-realists could be philosophical solipsists as well, but I take it this has little to say about the philosophy of science.

Up to now however, I think that the anti-realist really has no qualms about accepting our story. For, he will say, it licenses only the weak conclusion that we *might* find ourselves in a position where a claim for the existence of x-rays becomes worthy of belief.

The fundamental question is, why is there any ontological significance to the fact that, as a matter of fact and no more, we have not built (or met) these machines. Remember, science itself tells us that they are possible. As envisaged there is no doubt that we would accept them into our epistemic community. This in fact the anti-realist accepts, as a story. The anti-realist must claim that it is only because we do not *actually* interact with such machines that full belief in x-rays is not justified.

The anti-realist sees a bad argument right enough. This is the argument that could be ridiculed as follows: if we got good evidence that the earth would explode tomorrow then we would believe that the earth would explode tomorrow; we *might* get such evidence; therefore, we should *now* believe that the earth will explode tomorrow.

The argument involving hypothetical non-human rational observers is not the same. The sample argument just given should be recast as: if we got good evidence that the earth would explode tomorrow then we would believe that the earth would explode tomorrow; we know that if we encountered rational creatures of type X which we know to be possible they would sincerely inform us that the earth will explode tomorrow.

Now, what is the proper conclusion? Under these conditions, we would indeed have good grounds for believing that the earth will explode tomorrow.

The reason this second argument is compelling is that these creatures are possible members of the epistemic community. As possible members of the actual community they are, as it were, actual members of the possible ideal epistemic community. As such, their projected pronouncements are reports of the state of the world to which we owe credence *now*.

We already know that there *are* detectors of x-rays but the ones we are familiar with are dumb detectors. We admit that there could be smart detectors of x-rays as well, and if we met them we would then give full belief to the proposition that x-rays exist. So the only ground of ontological agnosticism here is the accidental impoverishment of our epistemic community, and this is no ground at all.

It seems to me that these considerations also tell against the second argument given above. That argument relied on the implausibility of the principle that our epistemic policies should give the same results independent of our beliefs about the range of evidence accessible to us. I would agree that we are not rationally compelled to accept such a principle, yet it seems we do not require such a strong principle to ground belief in unobservables. We require only something like the following principle: believe now what you would believe if you accepted as evidence what is unavailable *only* by accident. With respect to unobservables the accident is one of genetic background and evolutionary happenstance; it just happens that we cannot directly, that is, directly through the senses, have access to certain ranges of evidence. But this is just an accident which our theories can help us overcome. This is because it is our scientific theories which tell us what new members of our epistemic community would report and also because it is our theories which tell us what the range of beings capable of entering our epistemic community is.

The weak principle is surely much more plausible than the strong. Suppose you believe that if you looked in the cupboard you would find your keys. Assume only accidental events prevent you from looking--you're too busy, you have to go out right away, or whatever. Obviously you should nonetheless believe that your keys are in the cupboard. It is an admittedly more important and deeper accident that we have neither happened to have more extended senses nor either met or produced and accepted into our epistemic community any radically different sorts of intelligent creatures, but it *is* a sort of accident.

Refusal to accept this conclusion isolates the philosophy of science from the rest of the philosophical task of constructing a general picture of the world. Such a task obviously needs input from science itself. But it also requires the imagination to see how our epistemic practices must mesh with scientific information. If science aimed only at getting the relationships among observable phenomena correct it could not bear on the question of how we might integrate our view of the world with different views of the world, occasioned by different sensory contact with the world. Yet it obviously does bear on this question. But this suggests that, fundamentally, the only reason we could have for not believing in items unobservable to us is the accidental one of our place in the world. We could have had different senses; we could meet and accept as epistemic equals creatures with different senses. By the weak principle favored above, we then ought to believe in the unobservables postulated by science which such creatures could observe.

Notes

[1]We could reduce the level of idealization here by defining the notion of being in a 'deceptive epistemic situation.' The worst sort of deceptive situation is one where one is led to believe what is false, which we could very roughly define as: F is a property of being in a deceptive epistemic position if

$$(x)(Fx \rightarrow (\exists P)(B_x P \ \& \ -P)).$$

(A counterfactual clause is also probably necessary to the effect that if x was not F then x would not believe P.) P2 is then better stated by adding the rider that x does not believe that y is in a deceptive epistemic position.

[2]My choice of the word 'says' here is not meant to restrict in any very stringent way the modes of communication allowed between the members of the epistemic community.

References

Dennett, D. (1978). "Why the Law of Effect Will not Go Away." In *Brainstorms*. Bradford Books.

Maxwell, G. (1962). "The Ontological Status of Theoretical Entities" in *Minnesota Studies in the Philosophy of Science*, Volume 3. University of Minnesota Press.

van Fraassen, B. (1980). *The Scientific Image*. Oxford: Clarendon Press.

_ _ _ _ _ _ _ _ _. (1983). "Theory Comparison and Relevant Evidence." In *Testing Scientific Theories, Minnesota Studies in the Philosophy of Science*, Volume 10. Edited by Earman. University of Minnesota Press.