

P BUT I SHALL BELIEVE THAT NOT P

Pascal Engel

University of Caen and IUF

Colloque annuel de l'Institut Universitaire de France,

Nantes 1997

There is a well-known opposition, about the ethics of belief, between the "Victorian" position heralded by Clifford's famous maxim: "It is wrong, always, everywhere, and for everyone, to believe anything on insufficient evidence", and the so-called "pragmatist" view put forward by William James in his essay "The Will to Believe", according to which believing may be under the control of the will, and not strictly under the control of evidence. James argues that it is a matter of choice for us to avoid error:

"He who says: "Better go without belief forever than believing a lie!" merely shows his own preponderant private horror of becoming a dupe. He may be critical of many of his desires and fears, but this fear he slavishly obeys... A certain lightness of heart seems healthier than his excessive nervousness. At any rate, it seems the fittest thing for the empiricist philosopher."

The point is, presumably, not that we may believe falsehoods at will, or believe things which are contrary to our evidence, but that, in pursuing truth as a general

goal of our inquiry, we may have, as James says "a certain lightness of heart", and that we may believe things on insufficient evidence.

Bas Van Fraassen (1984, 1995) has recently defended a view about belief which, according to him, is in line with the Jamesian position, and which he calls voluntarism. Voluntarism in this sense is not the thesis that you may believe things at will, or decide to believe something at will, a view, which taken literally, faces obvious difficulties (Williams 1971). Voluntarism, for Van Fraassen, is rather the view that we are committed to our beliefs, and that we should be faithful to our beliefs. One way of expressing it may be through Moore' paradox :

P, but I do not believe that P

The Moorean statement is paradoxical, on this interpretation, because the person who utters it both commits herself to the truth of P by the first assertion and withdraws this commitment by the second assertion. This is because, according to Van Fraassen, statements of the form "I believe that P" ascribe a certain *property to a person*, that is a commitment to the truth of P, and not a certain *relation* between an individual and a proposition. Van Fraassen takes up a view put forward by Lewis (1979) according to which such statements are cases of self-attribution of properties. The property self attributed is the property of being a person *such that* one believes that P. Romantically, this could be expressed in this way: "I am, and I want to be, the sort of person who believes that P." Understood in this way, belief-predicates are not really descriptive predicates, predicates which describe states of an individual or which state a matter of fact, but are in a sense normative predicates, which assert that a person sets herself to believe that P. In *that* sense, according to Van Fraassen, belief is, to an extent (which has yet to be determined) a matter of the will, and not a matter of registering or describing one's cognitive attitude towards a proposition.

In the first part of this paper, I shall present Van Fraassen's argument for voluntarism. In the second, I examine the problems that it raises. I shall try to show that the Cliffordian view that Van Fraassen opposes is not really defeated by his version of voluntarism, and that it may well be compatible with it.

I

Van Fraassen argues for voluntarism about belief through the familiar framework of Bayesianism, according to which our beliefs and desires can be assigned degrees which are respectively degrees of subjective probability and degrees of desirability. According to Bayesianism in epistemology and in the philosophy of science, we form our beliefs on the basis of evidence, and we change our beliefs through conditionalisation adding new beliefs in our belief set conditional on our previous beliefs, according to the rule of conditional probabilities. Typically the Bayesian view can be called *past-oriented* : our new, *a posteriori* beliefs are dependent upon our past beliefs. Van Fraassen, on the contrary, wants to concentrate upon our future beliefs, and on how they can affect our current ones. If I discover that, within one year's time, I will believe that A, how should this affect my current attitude towards A ? In subjective probabilistic terms, if I have now a degree of confidence x in A, and then come to know that I shall have a degree of confidence y in A, where $y \leq x$, how should this affect my current degree of confidence? Van Fraassen answers that in general my degree of confidence x in A now, if I come to know that I will have a degree of confidence y in A in the future, should be *equal* to y , or lie within a range which is close to y , therefore that x should differ largely from y . If, I know that I will believe that A in, say, a year's time, to a degree between 0.7 and 0.9., then my present degree of belief in A should, optimally, be located between 0.7 and 0.9, or equal to the future assignation of degree. This is what he calls the *Principle of Reflection*, which he states thus:

To satisfy this principle, the agent's present subjective probability for proposition A, on the supposition that his subjective probability for this proposition will equal y at some later time, must equal this same number y .

$$PR = P_0(A / P_1(A) = y) = y$$

To illustrate this principle, suppose for instance that a weather forecaster announces today (Tuesday) that there is a 50% chance that it will rain on Friday, but also announces that he will, on Friday morning, predict that there is no more than a 25 % chance that it will rain that day. There is surely something odd with this weather forecaster, and the audience will be baffled. It will be baffled, presumably just as much as it would be baffled if the weather forecaster had said: "It will rain, but I believe that I will not rain". Well, maybe not quite, for we have allowed here that it is not a matter of "full belief", as Ramsey would say (belief with degree 1 or 0) in the occurrence of rain, but a matter of partial belief, and we know that such matters about our degree of confidence in the occurrence of weather conditions are rather vague. But even if we allow there that degrees of belief in the happening of rain are vague and likely to change, there is still an oddity in the weather forecaster's prediction, for it is, after all his job to assign precise quantities to his prediction.

Van Fraassen gives in "Belief and the will" an argument for showing that there will definitely be something wrong with such a weather forecaster. The argument uses the usual Bayesian strategy of making "Dutch books" to show that, we an agent is incoherent with his subjective probability attributions, a book may be made against him, which will have the result that he will loose money on a series of bets which a clever booker might make against him. Usual Dutch Book arguments, however, are typically synchronic: they consider only the agent present and past beliefs (and degrees thereof) and their coherence. Van

Fraassen's Dutch Book argument is diachronic, and considers not only the present and past beliefs but his future beliefs.

For the bizarre weather forecaster we have the following probability assignment (Christensen 1991):

$$(1) P_0 (A / P_1 (A) = 0.50) = 0.25$$

where "A" is the proposition that it will rain on Friday, $P_1 (A)$ the probability that it will rain on Friday as estimated today (Tuesday) by the agent, and

$P_0 (A / P_1 (A) = 0.50) = 0.25$ the agent's probability that it will rain on Friday given the agent's present estimation of the probability that it will rain on Friday.

we must also assume that the agent assigns some non zero probability to the proposition that on Friday morning he will think at a degree of 50% that it is likely to rain (if she did not, the conditional probability in (1) would be undefined). Let us suppose it is 20%:

$$(2) P_0 (P_1 (A) = 0.50) = 0.20$$

The Dutch bookie offers the agent two bets. The first bet will be won by the agent if tomorrow he believes, at degree 0.50, that it will rain. By (2), the agent believes that his chances to win the bet are 20%; so he puts 2 £ when the bookie offers 8 £, say. The second bet is more complicated. It is conditional on the agent winning the first bet (i.e conditional on his believing at degree 0.50 that it will rain on Friday). If this condition is not met, then neither the agent nor the bookie win anything. If it is met, then the agent wins the bet if it does not rain on Friday, and the bookie wins if it does rain. Thus his probability that he win the bet is 75%. So he considers it fair to put 30£ when the bookie offers 10£.

Now the bookie waits until Friday morning, when it is discovered whether the agent does believe at degree 0.50 that it will rain. If he does not so believe, then the second bet is null, and nobody wins. The bookie has, however, won the first bet, i.e he got 2£ from the agent. If the agent does believe, at degree 0.50, on Friday morning that it will rain in the Friday afternoon, then the agent has won

the first bet, i.e 8 £. And the second bet will not be null, but won or lost depending on whether it will rain or not in the afternoon. At this point the bookie offers the agent a third bet, which will be won by the agent if it rains in the afternoon, and by the bookie if it does not. Given that the agent assigns a probability of 50% to rain in the afternoon, he is willing to put 20 £ to the bookie's 20£.

The bookie can now be confident that he will win money anyway at the agent's expense. Suppose it rains. Then he has won the second bet, and he will win 30£. The agent wins 20£ on the second bet, and he has won 2£ on the first. So the bookie earns 2£. On the other hand, suppose it does not rain. Then the bookie has lost the second bet, losing the 10£ he put there, together with the 8£ he lost on the first bet. But he still wins the third bet, i.e 20£. So he ends up with 2£ profit here too. So the agent loses 2£ no matter what he believes on Friday morning and no matter whether it rains or not on the Friday afternoon.

He is incoherent in such a diachronic book, just as an agent is incoherent in a synchronic book. Van Fraassen claims that such situations validate the Principle of Reflection. He says:

"My integrity, qua judging agent, requires that, if I am presently asked to express my opinion about whether A will come true, on the supposition that I will think it likely tomorrow morning, I must stand my own cognitive *engagement* as much as I must stand by my own expressions of commitment of any sort. I can rationally and objectively discuss the possibility of discrepancy between my objective chances and my previsions. But I can no more say that I regard A as unlikely on the supposition that tomorrow morning I shall promise to bring it about that A. To do so would mean that I am no less than fully committed (a) to giving due regard to the felicity conditions of this act, or (b) to standing by the commitments I shall overtly enter." (1984: 255)

In this respect, believing — or rather the expression of one's belief— is like promising. If I express a promise which I also withdraw in the same breath, then I do not really promise.

II

The principle of reflection has some appeal. The case of the weather forecaster is obviously problematic, and the idea that we should stand behind our commitments is also quite plausible. But there are at least two *prima facie* problematic features with Van Fraassen's argument.

The first one is his reliance on Dutch book arguments and the Bayesian subjective probability framework. One difficulty with this framework is that we do not all the time assign precise degrees to our beliefs. A number of epistemologists in fact just deny that we reason or act out of partial beliefs, and claim that belief revision is made from full beliefs. In that respect one might say that we might not want to put odds on a certain proposition in accordance with our degrees of belief. So, if Van Fraassen wants to argue that belief implies a sort of commitment, why doesn't he argue for this point only through relying on Moore's paradox, which does not involve any subjective probabilities attached to our beliefs? And there is the familiar Pascalian point that even if we grant that the fear of betting losses gives us reasons not to have incoherent degrees of beliefs, the reasons that we have might not be epistemic. The reasons can be practical or utilitarian in the wide sense. According to Pascal's famous wager argument, the reason why one has to believe in God are not primarily epistemic reasons. They are reasons which have to do with the eternal bliss that we may gain from believing in God, the "infinity of gain", as Pascal says, which outweighs every evidential reason that we may have or not have for believing in His existence. In that respect, Van Fraassen's reliance on Dutch Book arguments is somewhat infortunate if we consider that he wants to defend a view like James'

in *The Will to Believe*, for James precisely makes appeal to Pascal's wager to claim that we value truth not for epistemic reasons only, but also because truth is connected with utility.

Van Fraassen, however, is well aware of these difficulties. This is why he does not rely too much, in a later paper (1995), on Dutch book arguments, and claims that they are rather of "heuristic value"(1995: 9). He rather appeals to Moore's paradox. The main problem with the weather forecaster in this respect, is not that he *contradicts* himself by believing that there is a 50% chance of rain while at the same time he believes that he will believe at degree that there is a 25% chance of rain. There is no *logical* inconsistency here. The problem is rather that by expressing his belief in this way, he undermines his credibility as an expert, just as I undermine my credibility when someone asks me: "Where is the philosophy department?" if I answer: "It's on the third floor, but I don't believe it's on the third floor." A more obvious way of pointing out that something has gone wrong with the weather forecaster is to compare the subjective probabilities with the objective ones. Subjectivists or personalists about probability do not claim that subjective probabilities can be assigned completely without taking into account the objective probabilities or frequencies of an event. It's not as if you could have any subjective probabilities you want, without caring about actual frequencies. Even subjectivists say that subjective probabilities should be *calibrated* with objective ones. In this respect, the weather forecaster's subjective probabilities go wrong because they cannot be so calibrated. Suppose we take a sample of 100 days with data identical with those that the forecaster has on Tuesday. Then the forecaster's Tuesday's prediction will be calibrated with the actual frequencies if it rains on half of those days, but the Friday morning prediction will be calibrated with the frequencies only if it rain only a quarter of those days. And it impossible for both predictions to be calibrated with the actual frequencies of the sample (Van Fraassen 1984: 245-246).

So the main problem is not so much the kind of incoherence that comes from Dutch books being made against the agent, and which show him irrational; the main problem is that the agent is not *diachronically* coherent, that is coherence between present belief and future belief. This is why Van Fraassen proposes in his 1995 paper not the Principle of Reflection above, but what he calls the *General Reflection Principle*, which runs like this:

GPR: My current opinion about event E must lie in the range spanned by the possible opinions I may come to have about E at later time E, as far as my present opinion is concerned.

And he says that this principle applies to someone who may only have *vague* degrees of belief about E. This why I remarked above that the weather forecaster case is odd, even if we agree that probability assignments to our beliefs are vague. Even if he says, on Tuesday, "There is *a chance* of rain Friday", and on Friday morning "It is *less likely* that I will believe that it will rain on Friday", he would still be guilty of diachronic incoherence. (Van Fraassen 1995: 15).

Let us suppose, for the moment, that the weakened principle of reflection GPR calms down the worries that we may have about PR. It would still not answer the following worry (Christensen 1991). There seems to be cases where the Principle of Reflection is violated and should be violated. Suppose that I take a certain drug which I know to have the following effect: it will, in a short while, make me confident that I can fly. But certainly it is not rational for me now to believe that I can fly. Or suppose that you are sure that you cannot drive safely after, say, ten drinks. Suppose also that after ten drinks, you would be sure (wrongly, as you now think) that you can drive safely. Reflection requires you to be sure that you would not be wrong if in the future you became sure that you can drive safely after ten drinks. But clearly the police would not advise to follow the Principle of Reflection here. These cases are similar to the well known story of Ulysses and

the sirens. Ulysses knows that hearing the Sirens sing will cause him and his sailors to pass overboard their ship into a dangerous sea. So he ties himself to the mast and puts wax in his ears, to protect himself and his crew from the temptation. Ulysses's case is a case of prudential step to avoid danger, and it has to do with practical rationality. It is sometimes a matter of prudence not to trust one's future beliefs. The voluntaristic advice about belief is a counterpart of this practical problem, but it gives us the opposite verdict. It says that we should not trust our present opinion when we know that it may be at odds with our future opinion. But the cases just described are cases where one should trust our present opinion, and not our future ones.

Perhaps one should try to restrict the scope of Reflection, and say that it applies only to the beliefs which are under my control. The counterexamples involve cases where there is a causal link between a certain state of my body and my future beliefs, which are not under my control. So we could try to rule out such cases of uncontrolled beliefs. Unfortunately the dividing line between what is under my control or not is not sharp.

The counterexamples to the Principle of Reflection are all cases where we know for sure that our future beliefs will be such or such, because there is some causal circumstance which more or less nomologically constraints these future beliefs, and which are such that our future beliefs will conflict with our present ones. But in general the boundary line between my present reliance in my belief abilities and my future reliance in my belief abilities is not precise. For instance, I know that, at the present moment, I can trust my memory; but I am not so sure that I shall be able to do so in the future. So if I can trust my future opinion about some event in the past, I shall be able to make this future opinion cohere with my present one about the same event. But if I do not trust my future opinion because I have reasons to think that I shall be unreliable, it is likely that there will be a conflict between my future opinion and my present one. So everything depends upon whether *now* I have reasons to believe that I shall believe this or that in the

future. I shall usually have some background information about my future reliability with respect to a given belief P. But if I don't, and in particular if the only reason I have to believe that P is one that is derived from a general attitude of trust towards my future opinions, then my future belief P give me only a weak reason to believe P now. And the reason is easily defeated when it conflicts with my current belief, as is shown by the counterexamples to the Principle of Reflection . (Foley 1994)

This means that both my current belief about P, and my current beliefs about whether I shall believe P in the future, have a sort of precedence over my future opinions, contrary to what voluntarism implies. Everything depends upon whether it is rational for me *now* to believe that P. It does not matter how you construe rationality or reasonableness of belief here. You can understand it in the "externalist" way, as a form of reliability in your present belief and belief forming capacities. Or you can understand it in the "internalist" way, as the presence of evidence available to you to confirm you beliefs. Whatever "rational" means here, rationality is tied to our present beliefs. It is only derivatively tied to coherence between past and future opinions. Once the agent is rational, it is incumbent upon him to make his present opinions cohere with future ones, but this coherence is not by itself a sign of rationality. In that respect the Principle of reflection is a maxim of rationality which is largely irrelevant to the formation of rational belief. To see this take the following case. I do not believe in God, for reasons which I take to be good: there is insufficient evidence for His existence. But I may be sure, or strongly confident, or just believe to a reasonable degree that at the time when I shall be approaching death, I shall believe in God for I know, or think I know, that it is common for people who are aware that they will be facing death in a not distant future to have religious feelings. Should I follow Reflection or not? On the one hand, one may argue that it irrational for me now to believe in God on the basis of the fact that I must assign the same degree to a belief in God now as the degree of the same belief in the future. So I should not

follow Reflection. On the other hand, it may be argued, along the voluntarist line, that it is quite rational for me to believe in God now, since it will cohere with my future belief. And I may be today closer to death than I think, so that my future belief does not lie in such a distant future. Of course the problem here is that it is not rational at all in the first place to believe in God because one fears that one is approaching death. Fear is not a proper ground for rational belief. But then it shows that the rationality or the irrationality of my beliefs have nothing to do with a diachronic coherence between my present and my future beliefs. It has to do with whether my beliefs, present, future, or past, are appropriately based on evidence. And if I think that they will not be based on further evidence when I shall be approaching death, then I should not trust my future beliefs, even when I know what they will be. But of course if my future beliefs can be based on appropriate evidence — if it were true, for instance, that when one faces death, one has more evidence available about the existence of God, say if it were the case that God sent to us more signs of his existence when we are closer to death than he does earlier—, then we should try to make our future (rational) beliefs cohere with our present beliefs, if the latter conflict with the former. So Reflection is supervenient upon the rationality of our beliefs, which is in general the rationality of our present belief, rather than the other way round.

If the exemple about my belief in God seems too moot, let us come back to the weather forecaster who believes that there are 50% chances of rain on Friday, but believes that on Friday morning he will believe that there are only 25% chances of rain. What goes wrong with the forecaster is that it is irrational for him to believe that he will have a different opinion on Friday, since that would show that he has not taken into account evidence for his Tuesday belief. It's not that he is irrational in not being coherent with what he will believe. the problem is rather that if he has reasons to believe that he will believe differently on Friday, then his *present* belief, that of Tuesday, must have relied on insufficient evidence. The

Dundee nov 95

incoherence between the future and the present belief is only due to the insufficiency of his evidence today.

So, after all Clifford was right: it is wrong, always, to base one's beliefs upon insufficient evidence. But James was right too in the following respect: once we have sufficient evidence for our future beliefs, and sufficient evidence that we shall have these beliefs, then we should try to adjust our present beliefs to our future ones. If this is what Van Fraassen mean by Reflection, it is a sound, and necessary, maxim of rationality. But it is hardly a sufficient one.