WILLIAMS ON NEGATIVE RESPONSIBILITY AND INTEGRITY

By John Harris

Bernard Williams's new essay, *A Critique of Utilitarianism*,[1] argues against consequentialism[2] largely by indicating ways in which a certain view of moral problems and dilemmas reveals the inadequacy of utilitarian solutions. His positive account of how moral problems are to be treated and of the ways in which crucial decisions are to be arrived at is what I wish to examine here, and since it is presented in the course of Williams's discussion of negative responsibility, it is on his treatment of this difficult problem that I shall concentrate. I should add that I have no wish to defend utilitarianism, the legions of the faithful will no doubt rally to its aid ; rather I seek to play Williams's part and attack another form of dogma which is fast taking over from utilitarianism its role as a piece of machinery for providing quick solutions to moral problems.

Integrity and Commitment

Williams introduces and explains the essential connexion between consequentialism and negative responsibility as follows:

> It is because consequentialism attaches value ultimately to states of affairs, and its concern is with what states of affairs the world contains, that it essentially involves the notion of *negative responsibility*: that if I am ever responsible for anything, then I must be just as much responsible for things that I allow or fail to prevent, as I am for things that I myself, in the more everyday restricted sense, bring about. Those things also must enter my deliberations as a responsible moral agent on the same footing. What matters is what states of affairs the world contains, and so what matters with respect to a given action is what comes about if it is done and what comes about if it is not done, and those are questions not intrinsically affected by the nature of the causal linkage, in particular by whether the outcome is partly produced by other agents.[3]

Consequentialism thus holds that there are two ways of making events occur in the world ; one is by positive actions, doing things with the result that these events occur, and the other is by negative actions, failing to do things with the result that these events occur. For those who wish to do so, the obvious ways to criticize the negative responsibility thesis are either to attack the causal linkage between inaction and consequence, and to claim that it is somehow more tenuous than that between action and consequence, or to " discredit it by insisting on the basic moral relevance of the distinction between action and inaction ". Williams, interestingly, chooses neither of these two methods. The latter he discounts as " unclear, both in itself and

[1] Bernard Williams, " A Critique of Utilitarianism ", in J. J. C. Smart and Bernard Williams, *Utilitarianism: For and Against* (Cambridge, 1973).

[2] Utilitarianism is, strictly speaking, a type of consequentialism, but for present purposes I use these terms interchangeably.

[3] *Op. cit.*, p. 95.

in its moral applications, and the unclarities are of a kind which precisely cause it to give way when, in very difficult cases, weight has to be put on it ". Neither does he attack the causal linkage (except at one crucial point which I shall examine in due course) but rather assumes it. Instead he deploys an argument designed either to justify our causing harm by our negative actions in certain circumstances, or at least to defend our right to do so in these circumstances. It is this argument that I shall now examine.

THE ARGUMENT FROM INTEGRITY

The kingpin of the argument that Williams deploys against the negative responsibility thesis is the special and necessary value that each person must place upon his own integrity. For Williams a person's integrity is bound up with a class of projects to which the individual is particularly *committed*.

> How can a man as a utilitarian agent come to regard as one satis-
> faction among others, and a dispensable one, a project or attitude
> round which he has built his life, just because someone else's projects
> have so structured the causal scene that that is how the utilitarian
> sum comes out ? The point is that he is identified with his actions
> as flowing from projects and attitudes which in some cases he takes
> seriously at the deepest level, as what his life is about. . . It is
> absurd to demand of such a man, when the sums come in from the
> utility network which the projects of others have in part determined,
> that he should just step aside from his own project and decision and
> acknowledge the decision which utilitarian calculation requires. It
> is to alienate him in a real sense from his actions and the source of
> his action in his own convictions.[4]

Williams takes for granted our nodding assent that indeed a man *cannot* step aside from his own projects to obey a utility machine, which has been programmed by the projects of others. He reinforces his claim that a man, in some sense of the word, *cannot* dispense with his commitments, by suggesting that he has a right to defend his integrity, and by demonstrating that the categorical imperatives of utilitarianism constitute, in some circumstances, a fundamental attack on this integrity. We must concede at once that it would be absurd to demand of a man that he sell his soul to the utility machine, binding himself to throw over his own projects and deeply held attitudes wherever and whenever some slight gain for utility might thereby be achieved. And it might even be true that in some sense a man literally cannot, in the comprehensive way that utility demands, abdicate the direction of his own life and still keep the autonomy and integrity that give his life coherence, that make him an individual. But now how much of the thesis of negative responsibility have Williams's arguments demolished?

The first thing to note is that his weaponry, like the H-Bomb, does not enquire too closely into the identity of its victims. He has produced a general argument to show that maintenance of integrity necessarily involves a certain degree of insensitivity to the needs or wants or to the happiness of others. But this attacks positive quite as much as negative responsibility. If it is absurd to demand of a man that he abandon his projects because someone has so structured the causal scene that a greater utility lies in his doing something else, then it is equally absurd to demand of him that he abandon a project which directly involves the same amount of disutility. Remember, we are talking about projects and attitudes that a man " takes

[4]*Op. cit.*, p. 116.

seriously at the deepest level as what his life is about "; it would be as absurd to demand that a man abandon projects as important as *these* if some fractional general disutility were directly involved, as it would be if an equally small general disutility stemmed simply from his failing to do something else. If Williams wishes to demolish negative responsibility alone he must undermine its foundations, either by attacking the causal linkage, or by defending some sort of moral difference thesis, neither of which he in fact does, with the exception already noted but yet to be discussed.

While it may be true that a man cannot sell his soul wholesale to the utility network, it is not true that in any particular case he cannot sacrifice his projects to utility. To put this point in the language of negative responsibility, a man *can* come to regard his deeply held projects as dispensable when circumstances, even circumstances determined by others, force on him, or provide him with, the opportunity to prevent serious harm. So that in any case worth discussing it will not be true that a man cannot, in any sense of the word, choose to abandon his projects. What he will have to do is choose whether to abandon them or not.

So we are left with the weaker position which argues not that we *cannot* abandon our projects when to do so would prevent harm, but that we need not do so. Now if this means that we need not necessarily or automatically abandon our projects whenever by our doing so a net increase in general utility could be obtained, then again we can concede the point. For again to do so would be to abdicate the direction of our lives to an extent that would rob us of the autonomy and integrity that give our lives their coherence and individuality. We must be clear that even here Williams's argument does not cut off negative responsibility. His argument does not show that we are not responsible for the harm that we fail to prevent, only that there are cases in which we are justified in not preventing it. The crucial questions are: which are these cases, and how are we to decide which they are ? And again, in any cases in which the harm we might prevent is sufficiently serious to make us even so much as think about whether to prevent it or not, then we must surely ask ourselves whether there are considerations of sufficient moral weight to justify our not trying to prevent it. But it is just this sort of question which it seems that Williams wishes to protect us from having to ask. The special and necessary moral weight which he attaches to our integrity is supposed to relieve us from the necessity of balancing projects with which we are " more deeply and extensively involved " against the harm which by abandoning them we could prevent. But how can anyone even remotely interested in doing the right thing in a given situation avoid this sort of calculation ?

To recapitulate: negative responsibility holds that I am just as responsible for things that I allow or fail to prevent as I am for things that I myself, in the more everyday restricted sense, bring about. Nothing that Williams says damages this thesis. At most his arguments show that there are cases, perhaps many cases, where we are justified in, or can be forgiven for, occasioning some harm to others. But this is not news, and it applies as much to harm for which we are positively responsible as it does to harm for which we are negatively responsible ; or if it does not, nothing Williams says shows that it does not.

As an attack on utilitarianism, Williams's argument from integrity has some success. He is persuasive in claiming that we cannot or need not *automatically* abandon our deeply held projects *whenever* any increase in general utility could thereby be obtained. But this leaves all the crucial

questions unanswered. What we need to know is how far his argument
from integrity is supposed to justify our being responsible for harm to others
in any particular case. If a man is committed to his work or to his wife,
can he while sitting at his desk involved in some vital calculation or by her
side engaged in a crucial quarrel ignore the cries of a child drowning in the
pond outside ? Williams goes no way towards persuading us that our integ-
rity is of such overriding importance as to justify our not bothering to decide
in any particular case whether our project is sufficiently important for us
to pursue it at the cost of the harm which by abandoning it we could pre-
vent, and again, this is just the sort of calculation that Williams wants to
short-circuit. Irritatingly, when Williams gets down to cases he does not
employ his argument from integrity at all, but rather offers some very differ-
ent suggestions.

NEGATIVE RESPONSIBILITY: WILLIAMS'S EXAMPLES

Williams provides two fascinating and complex cases, but their fascina-
tion and complexity render the prospect of drawing any clear conclusions
somewhat shady. It is clear that Williams values these cases primarily as
counter-examples to utilitarianism ; and they may well be, but one at least
is also a counter-example to Williams's own argument from integrity. Wil-
liams seems to realize this, for he produces a new suggestion to accommodate
this case, a suggestion moreover that puts morality beyond the scope of
rational debate. These are the cases:

(1) George, a research chemist with a wife and small children, cannot
find a job. A friend can procure for him a job at a chemical and biological
warfare establishment. The job is not to George's taste but if George does
not take the job a zealous contemporary will. What should George do ?

(2) Jim finds himself in the central square of a South American town.
Twenty Indians are about to be executed, but the captain of the soldiers
(call him Pedro) offers Jim this alternative: either Jim can kill one of the
Indians with his own hand and the others will be allowed to go free, or, if
he refuses, all twenty will be shot. What should Jim do ?[5]

In George's case we are to judge that the utilitarian answer is that he
should take the job, thus providing for his family and enabling him to slow
down or foul up research, maximizing utility all round. The alternative,
sticking to his principles and preserving his integrity, leaves himself out of
work, his family unprovided for, and an unprincipled maniac pushing ahead
with inhuman research. In the second case the utilitarian answer is more
obvious. Jim should kill one man to save nineteen. Jim's dilemma stems from
his unwillingness to kill anyone and from his distaste for being dragged willy-
nilly into an unsavoury situation of someone else's making, in which he
must choose between alternatives equally repugnant to him.

Williams concludes that it is not hard to see that in George's case, viewed
from the perspective of integrity, the utilitarian answer would be wrong.
But this is, I think, because what Williams takes to be the utilitarian answer
is not really the utilitarian answer. But let us grant that utility might just
be on the side of George's taking the job. There is his family to consider,
and we can, I suppose, imagine that it might on balance be safer for man-
kind, and so consistent with utility, if, as a precaution, men like George
worked in places like chemical and biological warfare establishments. And

[5]*Op. cit.*, pp. 97 ff. In the original Pedro is the captain's hatchet-man, but it simpli-
fies the argument to ignore this unnecessary multiplication of murderers.

in such a case I think Williams is right in supposing that most people would agree that George need not (i.e., it is not the case that he ought to) take the job ; after all, the harm that George, by taking the job, *might* prevent is itself highly speculative.

But now compare Jim's case. Nineteen or twenty lives are definitely and immediately at risk and the action required to save nineteen of them is not the work of a lifetime but of a second or two. Williams clearly senses the difference between the cases for he thinks that the utilitarian judgement about Jim's case is probably the right one.

> The immediate point of all this is to draw one particular contrast with utilitarianism: that to reach a grounded decision in such a case should not be regarded as a matter of just discounting one's reactions, impulses and deeply held projects in the face of the pattern of utilities, nor yet merely adding them in—but in the first instance of trying to understand them.[6]

Here again I agree with Williams, but he does not appear to agree with himself. One would expect him to go on to give an account of how Jim's integrity, his various commitments, and his deeply held projects and beliefs might weigh with him in coming to a grounded decision ; and of whether the sort of weight these considerations might have would be sufficient to counter-balance the nineteen lives in the scales with them. But Williams does not do this, for the very good reason that one cannot imagine what commitments Jim might have that would warrant the sacrifice of nineteen lives. What Williams does is throw out a number of suggestions that he believes Jim's case and others involving negative responsibility will " have to take seriously ". But before looking closely at these suggestions we must, as promised, consider Williams's attempt to attack the causal connexion between inaction and consequence. It comes in the context of Jim's case.

> For Pedro's killing the Indians to be the outcome of Jim's refusal, it only has to be causally true that if Jim had not refused, Pedro would not have done it.
>
> That may be enough for us to speak, in some sense, of Jim's responsibility for that outcome, if it occurs ; but it is certainly not enough, it is worth noticing, for us to speak of Jim's *making* those things happen. For granted this way of their coming about, he could have made them happen only by making Pedro shoot . . . it is misleading to think in such a case of Jim having an *effect* on the world through the medium (as it happens) of Pedro's acts ; for this is to leave Pedro out of the picture in his essential role of one who has intentions and projects, projects for realizing which Jim's refusal would leave an opportunity.[7]

But if it is misleading to think in such a case of Jim's having an *effect* on the world through the medium of Pedro's acts, because this leaves Pedro out of the picture in his essential role as someone capable of forming the intention not to shoot, then it is equally misleading to think of Pedro having an *effect* on the world, for this leaves Jim out of the picture in his essential role as someone capable of assenting to Pedro's proposals and thus (as it happens) preventing Pedro from shooting. Williams's way of setting up the case stipulates the truth of the counterfactual " If Jim had not refused,

[6]*Op. cit.*, p. 118.
[7]*Op. cit.*, p. 108-9.

270 JOHN HARRIS

Pedro would not have done it ", so both Jim and Pedro have an essential role to play. If twenty are to die the " co-operation " of both men is necessary, and if one is to die their " co-operation " is necessary ; that is the way the example is set up. Of course what Williams is after is a difference in the causal efficacy of Pedro's and Jim's respective contributions to the deaths of twenty Indians. But this is just what he cannot have ; for while Jim cannot make Pedro shoot, Pedro cannot make Jim assent, and if Jim assents, Pedro can no more shoot than he can if (as it might happen) he decides not to. For *this* is Williams's case, and if it is changed so that Pedro is going to shoot whatever Jim does, then it is not of course even causally true that if Jim had not refused Pedro would not have shot.

Williams has put all the weight of his argument on the very tricky notion of *making* things happen. It is a general feature of negative acts that the agent does not make things happen but rather allows them to occur, but it is far from being a general feature of positive actions that the agent *makes* things happen, at least in the sense that Williams seems to be relying on. I take this sense to involve some idea of the agent's being in complete *control* with the capacity to *ensure* that the event will happen. If I shoot a man intending murder he may die only if my bullet is not deflected, and moreover hits him in a vital spot, and perhaps only then if the victim does not get to a doctor in time, or if the doctor happens to be inefficient. I cannot *make* him die, I can only do my best ; but if he does die, the fact that in shooting I did not positively *guarantee* his demise will not be taken as a mitigating factor. With positive acts it will always be a question of fact whether or not the act does indeed have the expected or probable consequences, and the same is true of negative acts. Part of what Williams seems to be relying on in Jim's case is the fact that Jim is powerless and Pedro literally in control. Pedro can make sure that he kills all the Indians if he wants to and Jim can do nothing to stop him but must rely on Pedro's word. But if we change the example so that Jim is in hiding with a machine-gun in the use of which he is an expert, then Pedro cannot make sure of the Indians, he only supposes that he can. For Jim can cut him down even as his finger whitens on the trigger. In this case he is not as might appear the most proximate cause of the deaths of the Indians (if they die) for he needs the acquiescence of Jim, he can only make things happen if Jim lets him. While it is true that Jim cannot make him shoot, nothing Williams says shows that this is a morally or causally relevant fact, any more than it would be morally or causally relevant to say that a doctor cannot *make* a man die of loss of blood by withholding a necessary transfusion.

OLFACTORY MORAL PHILOSOPHY

We must now return to Williams's serious suggestions for dealing with negative responsibility. One of these is the alarming suggestion that a grounded decision " might not even be decent ":

> Instead of thinking in a rational and systematic way either about utilities or about the value of human life, the relevance of the people at risk being present and so forth, the presence of the people at risk may just have its effect. The significance of the immediate should not be underestimated . . . very often, we just act, as a possibly confused result of the situation in which we are engaged. That, I suspect, is very often an exceedingly good thing.[8]

[8]*Op. cit.*, pp. 118.

It is interesting to compare Williams's attitude here to Tolstoy's. Talking of Levin towards the end of *Anna Karenina*, Tolstoy remarks:

> Whether he was acting rightly or wrongly he did not know—indeed, far from laying down the law, he now avoided talking or thinking about it.
>
> Deliberation led to doubts and prevented him from seeing what he ought and ought not to do. But when he did not think, but just lived, he never ceased to be aware of the presence in his soul of an infallible judge who decided which of two courses of action was the better and which the worse, and instantly let him know if he did what he should not.[9]

Williams seems to be recommending here the use of what Orwell, following Nietzsche, called " moral nose ".[10] The nose of the basically decent man, being a well adjusted instrument, will tell him what to do, and he will act for the best. But for this sensitive source of moral insmell to operate it is necessary not only for the immediate to be significant, but for us to be confronted with it, to be as it were within sniffing distance. If we are in the presence of the people at risk, their presence might well have its effect and this effect might be for the best. In as much as Williams regards the immediate as a stimulus triggering an automatic response, that is all that can be said, for this response either will or will not occur and either will or will not be for the best. But there is also the suggestion that a grounded decision " might not even be decent ". Consider another recent defender of what we might call the olfactory school of moral philosophy. Noam Chomsky feels that " by entering into the arena of argument and counter-argument, of technical feasibility and tactics, of footnotes and citations, one has already lost one's humanity ".[11]

But this insistence on the moral priority of the nose is disturbing. For much if not most of what should concern us morally takes place beyond the limited range of our organs of moral sense. If we are to act for the best we must ground not only particular decisions, but the conduct of our lives, on a careful consideration of the many different features Williams points to, seemingly only to reject. " Out of sight " must not become the justification not only for " out of mind " but also " out of account ". Life is no longer easy for the olfactory moral philosopher. " Civilization " has been defined as a device for shielding mankind from a cross-section of human experience. The poor are often hidden away in slums, the sick or dying in hospitals, the eccentric or depressed in asylums, the aged are left to die of malnutrition or bronchitis or of cold in the privacy of their own homes, and famine victims live in foreign countries.[12] Moral nose cannot be relied on to prompt us to action on behalf of these people.

The other obvious drawback of moral nose is that we shall want, or we ought to want, to know whether our response to the immediate is the right

---

[9]Leo Tolstoy, *Anna Karenina*, trans. Rosemary Edmonds (Harmondsworth, 1954), p. 826.

[10]George Orwell, letter to Humphry House, 11 April 1940, in *The Collected Essays Journalism and Letters of George Orwell*, Vol. I, (Harmondsworth, 1970), p. 583.

[11]Noam Chomsky, *American Power and the New Mandarins* (Harmondsworth, 1969), p. 11.

[12]It is one of the ironies of civilization that having swept the unfortunate out of sight we have to be made aware of their existence, and our sympathy aroused, through advertising.

one ; and this we can only find out by trying as best we can in the perhaps limited time available to weigh all the relevant considerations and come to a grounded judgement.

OTHER PEOPLE'S PROJECTS AND OTHER PEOPLE'S NEEDS

Another consideration to which Williams attaches great importance, and which he introduces a number of times, is the distinction between my projects and other people's projects. It is from the perspective of other people's projects, as we have seen, that Williams wishes to criticize negative responsibility.

> Discussions of [Jim's case] will have to take seriously the distinction between my killing someone, and its coming about because of what I do that someone else kills them: a distinction based not so much on the distinction between action and inaction, as on the distinction between my projects and someone else's projects.[13]

It is not clear why Williams thinks this a useful distinction. At best it could only touch a small corner of negative responsibility, for the simple and sufficient reason that most of the harm for which we may be negatively responsible is not attributable to the machinations of any men besides those who decline to prevent it. It does not arise from the projects of other people, but rather from such " natural " causes as disease, famine, drought, etc. The only people involved are those at risk and those who can help them. The weight of Williams's attack on negative responsibility must be borne by the argument from integrity and in this context other people's projects would figure merely as a class of need-creating circumstances, significant only because they highlight the centrality of our own integrity by creating needs which, in crying out to us for help, interfere with our projects.

But in those cases of negative responsibility which do involve other people's projects, and it is significant that both his examples are of this sort, Williams has a special argument. His handling of these two cases suggests that he is really interested in the autonomy and independence of individuals.[14] His point is that if George and Jim do as utility seems to dictate, they would also in effect be surrendering to an alien intelligence, allowing their decisions to be virtually dictated by the projects of another person. This is perhaps why in Jim's case, where his conclusion is that Jim ought to kill one Indian to save nineteen as Pedro suggests, Williams wants Jim to act automatically, to respond directly to the distress of the Indians and not go through the calculations imposed on him by the way Pedro has structured the situation. By following what Tolstoy referred to as " the infallible judgements of [the] soul ", Jim can avoid making any of the calculations Pedro would like to force upon him, he can avoid accepting the range of alternatives that have been imposed by another. In effect he ignores the other person and his projects, and by consulting only his own soul he avoids becoming merely a pawn in another man's game, and remains his own man.

But this shows that the distinction between my projects and other people's projects must be a red herring. Williams wants Jim to come to the same conclusion that he would have come to if he accepted the range of alternatives imposed on the situation by Pedro, but he does not want him to come to the conclusion *that way*. Jim must act without thinking so that he can

[13]Williams, p. 117.

[14]See for example Bernard Williams, *The Idea of Equality*, in Peter Laslett and W. G. Runciman, edd., *Philosophy, Politics and Society* (Oxford, 1962), p. 114.

both preserve his integrity *and* come to the right decision. But this sort of attempt to have one's moral cake and eat it is likely to lead to a nasty attack of *mauvaise foi*. Either people must try to convince themselves that they do not know why their actions are right or they are condemned never to wonder whether they have in fact done the right thing. Other people's projects must, after all, be seen simply as a class of need-creating circumstances and must, for the purposes of moral decision-making in cases such as George's and Jim's, be treated simply as features of a brutal world, like famine and disease. What I have to choose between is not my projects and other people's projects but my projects and other people's needs. As a moral agent I must weigh what they will suffer if I do not help them against what I will suffer if I do. And whether or not in bringing others help I am forced to choose between alternatives imposed on the situation by others cannot make a ha'p'orth of moral difference to me. If I believe that other people are going to carry through their projects with harmful consequences which I can forestall, then for the purposes of moral decision-making I must treat their projects simply as need-creating circumstances. The fact that these people can relent is not relevant if they are not in fact going to relent. The bare fact that it is possible for them to relent can in no way absolve me from the moral duty of weighing the extent of the harm they will cause against the importance of the project I must abandon to forestall the harm.

CONCLUSION

As a way of criticizing negative responsibility Williams's argument from integrity is a non-starter.[15] As we have seen, it applies as much to positive as it does to negative responsibility and in neither case does it cut off our responsibility. Integrity is a factor which might serve to justify our being responsible (positively or negatively) for the harm in question. But it does not help us with moral dilemmas, for the dilemmas, at least the dilemmas Williams considers, are about what price to put on our integrity. Despite disclaimers, Williams talks as though once we realize that our integrity is involved we no longer have to weigh at its full value the cost of persisting in our projects ; but he offers no argument to show why this should be so in any particular case.

His olfactory recommendations are offered without apology as a way of solving and not merely of looking at moral dilemmas. And here we must conclude that far from being " very often an exceedingly good thing " to solve moral problems nasally, it can never be right to do so. For if Williams were right about olfactory moral philosophy it would be in effect just like the utility machine—simply a piece of apparatus designed to yield instant solutions to complex and agonizing moral problems.

Balliol College, Oxford

---

[15]For a more detailed discussion of some attempts to circumscribe or deny negative responsibility, see my papers " The Marxist Conception of Violence " (*Philosophy and Public Affairs*, Winter 1974), and " The Survival Lottery ", forthcoming in *Philosophy*.