

ON THE FUNCTION OF FALSE HYPOTHESES IN ETHICS.

C. D. BROAD.

THE title of this paper is obscure, but the question which it proposes to discuss is important; it is closely connected with an ethical principle which is perhaps more often explicitly used in the reasonings of daily life than any other. This principle or mode of argument I call the *Principle of False Universalisation*.

A man proposes to himself a certain course of action and debates whether it be right or wrong. At a certain stage he will say to himself, or, if he be discussing the matter with a friend, his friend will say: Suppose *everybody* did what you propose to do. The consequences of this hypothesis will then be considered, and, if they be found to be bad, the man will generally consider that this fact tends to prove that his proposed action is wrong. I think the principle is nearly always used negatively, *i.e.*, to condemn a proposed course of action. We do not in general argue that a proposed action must be right because if everybody did likewise the result would be excellent. How far there is anything to be said for this distinction we may see in the course of our discussion.

Everybody is familiar with this kind of reasoning; everybody seems to think that it is valid and important; and, as we know, something very like it was regarded by Kant as *the* fundamental law of ethics. Yet people have not noticed how extremely paradoxical it is, and what curious and interesting questions it raises on the border line between ethics and logic. The paradox is this. We are asked to believe that the rightness or wrongness of many of our actions depends on the probable consequences, not of what we judge to be true, but of what we know to be false. For, in practically every case where we consider what would happen if everybody acted as we propose to

act, we know as surely as we can know anything that is not *a priori*, that by no means everybody will act this way. *E.g.*, a man says: I should intensely dislike to be married and don't intend to marry. His friend replies: But suppose everybody refused to marry, would not the results be very bad? If the man agrees that the results would be very bad he will be inclined to think that this is *pro tanto* an argument against the rightness of his refusal to marry. Yet both he and his friend are perfectly well aware that the hypothesis which they are supposing is false; they know quite well that as a matter of fact there is not the least prospect of everybody refusing to marry or even of so many people remaining unmarried as to lead to the consequences which they agree would be bad if they actually existed.

It might of course be said that the paradox only arises when you consider that the rightness or wrongness of actions depends on their actual or probable consequences, and that it is avoided if you suppose that it depends wholly on their own intrinsic qualities. So long as we believe that probable consequences are relevant in deciding the rightness or wrongness of an action, the particular circumstances under which the action is to be performed must be taken into account, since its probable consequence will largely be determined by them. And a very important circumstance must be the question whether other people are or are not going to do similar actions. Hence, if probable consequences are to be considered at all, we cannot and ought not to be guided by a false account of the circumstances; and the hypothesis whose consequence we are asked to consider in the method of false universalisation is admittedly a false account of the circumstances in which our proposed action would take place. If we had only to consider the intrinsic quality of the action this difficulty would vanish. For, since we are not to consider circumstances at all, the mere fact that the universalisation gives a false account of them need not trouble us. But, as against this, two things must be said: (1) It is thoroughly

unreasonable to suppose that the goodness or badness of an action is entirely independent of its probable consequences, and no one but a moralist riding a theory to death would maintain this view for an instant. (2) Such a view can hardly be held consistently by persons who support the method of false universalisation. For they are admittedly asking us to consider and appraise consequences; and they can hardly take up the extraordinary position that it is only the consequences of what we know to be false that are ethically relevant, while the consequences of what we believe to be true are wholly unimportant.

And, as a matter of fact, except in the case of false universalisation, it would commonly be held that it is only the probable consequence of what is believed to be true that is ethically relevant. To be guided by the probable consequences of what is known to be false would, in most cases, be regarded as absurd and immoral. A man who proposed to establish a system of communism would not be held to have justified himself if he merely said: Consider how excellent the consequences of my system would be if no one were guided by selfish motives. Yet the hypothesis that no one is guided by selfish motives is not more obviously false than the hypothesis that no one is prepared to marry. But, seemingly, the admitted falsity of the former hypothesis makes its consequences ethically irrelevant, while the admitted falsity of the latter does not. There is a real paradox here, and it leads us to the general question: Is the goodness or badness of the consequences of admittedly false hypotheses ever relevant in deciding rightness or wrongness of a course of action; and, if so, what is the distinction between those false hypotheses whose probable consequences are relevant and those whose probable consequences are not?

We may usefully begin by considering cases where the employment of false hypotheses is obviously justifiable, and gradually working up to the more difficult and paradoxical examples of its use which we have just been illustrating. The first and simplest use is for the avoidance of personal

prejudice. We feel inclined to perform a certain action and are not sure whether it is right. We then say: Suppose Smith were in my circumstances and did the action that I propose to do, what should I think of him? If, when we try to envisage this false hypothesis, we find that we should strongly condemn it, we may be fairly sure that our own proposed action is wrong and that our tendency to approve it in ourselves is due to mere personal prejudice. This obviously justifiable and extremely useful employment of false hypothesis raises no theoretical difficulties. We are not supposing that our act takes place under circumstances different from the actual in any relevant respect; we are supposing that its consequences are the same and that the only difference in circumstances is that it is performed by Smith and not by me. And it is essential that no other differences should be introduced. If I introduce considerations about Smith's particular situation, or consider any difference between him and me other than the fact that he differs numerically from me, my judgment based on contemplating his action can no longer be applied straightway to mine. And the reason is that the action will now have been performed under different circumstances and will therefore probably have different consequences from my own proposed act. If the question is: Ought I to buy a motor-car? and I find nothing to disapprove when I contemplate the hypothesis of Smith buying a motor-car, this will have no bearing on the rightness of my proposed action if one of the differences that I have taken into consideration is that Smith's income is twice mine. The essential limitations then of this use of false hypotheses are such that it throws no light on the question of false universalisation; for it differs just in the circumstances which make false universalisation so paradoxical.

Another use of false hypothesis is where we decide on the right course of action by considering what somebody whose moral character we admire would probably do if placed in our position, *e.g.*, some evangelical Christians

claim to solve moral problems by considering: "What would Jesus do?" No special logical difficulty is involved here, and the method is simply of the following type: I cannot see by direct inspection or by reasoning what is the right course of action here. I can be sure that if Jesus were in my situation His action would be the right one, and I am well enough acquainted with the character of Jesus to be tolerably certain what course of action He would take. I shall therefore be safe in following that course of action. Whatever we may think of the practicability of such a method, there is nothing recondite in it; it is of the same logical type as looking up a recipe in cookery book fortified by the knowledge that Mrs. Beeton was a better cook than I am. We are not, as in the method of false universalisation, considering the results of an action done under circumstances that we know not to exist. We suppose that Jesus is in exactly the same position as we are and make abstraction of all differences between Him and us except the difference that He knows what is right and will certainly do it, while we are weak in insight and in performance. If we take into account other differences, we cannot reasonably argue from what Jesus would do to the rightness of a similar action on our part. A being who could raise the dead would *e.g.* be justified in performing certain operations for which a human doctor could justly be blamed. It is perhaps worth while to notice that these considerations wreck the method in practice. To obtain useful ethical conclusions about our own actions, we need to regard Jesus simply as differing from us in insight and good-will. But, to obtain conclusions as to His probable course of action in given circumstances, we must argue inductively from what we know about His recorded actions; and these are the actions of a being differing from us in innumerable other respects beside the two mentioned.

These two examples of false hypothesis then throw very little light on the particular problem that interests us, viz., that of false universalisation. But we can now pass

to certain genuine cases of false universalisation where the logical principles involved are easy to recognize. We will discuss these before passing to the most difficult and paradoxical uses of this principle. The use that we have now to consider may be described as the use of the principle of false generalisation as a moral microscope.

The result of one man's action may be very small, and it may be impossible for him to see by contemplating it alone whether it be good, bad, or indifferent. But he may be able to see that a great number of such actions would produce a result of the same kind as a single one but of much greater magnitude, and that this result would be unmistakably good or bad. If he has reason to suppose that the goodness or badness of the results of a large number of similar actions is the sum of the goodness or badness of the results of the separate actions, he will be able to conclude as to the moral quality of his own proposed action though it was not obvious on mere inspection. It is clear that if such an argument be ever applicable, the falsity of the hypothesis is irrelevant. We are admittedly capable of estimating the goodness or badness of merely hypothetical states of affairs. This being granted, the general line of argument runs as follows: If the complex *C* existed it would be good (or bad). Now the complex *C* contains a part *c* precisely similar to the results of a certain contemplated action. We have reason to believe that *C* could not be good (or bad) unless *c* were itself good (or bad). Hence we can come to a conclusion about the moral quality of our proposed action even though this be too small to reveal itself to mere inspection.

The nerve of the argument obviously is the condition that a complex shall be known to be so related to one of its parts that the former cannot be good (or bad) unless the latter has the corresponding moral predicate. Unfortunately, the principle of organic unities makes it extremely difficult to be certain in any particular case that this relation subsists between the value of a whole and that of one of its parts. It seems most plausible to suppose

that this relation holds for those complexes which consist of a number of precisely similar parts. And these are the cases contemplated when the principle of false universalisation is used as a moral microscope. The appeal to everybody is here, strictly speaking, rhetorical; all that is really necessary is to consider the results of a fairly large number of people performing actions precisely like the one under discussion. And the argument, if *ever* valid, is subject to great limitation and doubt. We must take the results of our act quite in abstraction before hypothetically multiplying them, and this is liable to be overlooked if we bring in the notion of other agents performing precisely similar acts. To take a typical instance. I walk through a field and pluck an ear of corn. Is this right, wrong or indifferent? If I now say: Suppose a million people walked through and each plucked an ear, the results would be very bad, I must of course make abstraction of the effects of a million people merely walking through. My walking through may have done no damage whatever, but it would be physically impossible for a million people to walk through without doing grave damage. It is better, therefore, not merely to drop the reference to everybody, but also to drop the reference to agents altogether and to consider nothing but the hypothetical multiplication of results like the result of my action.

But, even when this is done, it seems to me that the argument from the damage done by a million ears being plucked to that done by the plucking of one is most precarious. The consequences that have to be considered cannot be the mere separation of the ears from the stalk; this, like all physical events, is in itself morally indifferent. We obviously have to go further and consider the effects on the state of mind of the owner of the field and of others. Now it seems perfectly possible that no one's state of mind is in the least better or worse for the plucking of one ear and yet that it may be very much the worse for the plucking of a million. There is absolutely no logical reason against this and it seems to me to be true. The most

probable account of the matter is that the plucking of a certain finite number n (varying of course with the circumstances) is absolutely indifferent, while the plucking of any greater number leads to consequences which get worse as the number gets greater. It is no objection to this view that we cannot state exactly what the number n is; for it is no objection to any theory that it does not presuppose omniscience in its supporters. I think we may put the objection to the moral microscope as follows: If you merely consider physical consequences, they have no moral value and therefore cannot help you to decide any moral question. If you go on to consider psychical consequences, then there seems hardly any reason to believe that the psychical consequences of a large number of precisely similar physical events form a whole consisting of a number of similar parts and itself similar to its parts, or a state differing *only* in magnitude from that produced by each physical event separately. And it is not at all clear to me that if two psychical states differ only in their intensive magnitude and the one with the greater intensive magnitude be good (or bad) the one with the less intensive magnitude must have the same ethical quality but to a smaller degree.

I conclude then that both on practical and ethical grounds it is most unlikely that you can ever safely argue from the goodness or badness of the effect of a number of precisely similar acts to the rightness or wrongness of a single act of the class. And, unless this can be done, the moral microscope can have no valid practical use.

But we now come to a class of cases where the principle is used, but where the moral microscope explanation, even if it be valid, can hardly be the full explanation. And these are just the most paradoxical and perhaps the commonest instances of the use of the principle of false universalisation. The cases that we have to consider are the following: A certain good can only be produced by the co-operation of a number of people. The acts of these people need not, of course, and in general will not be *precisely* similar; all that is important is that they resemble

each other in the fact of co-operating to this end. The use of the principle that we have now to consider is to prove to people that they ought to co-operate. A man is supposed to admit that the results obtained by the action of a certain group are good, and that they can only be obtained by co-operation. He refuses, however, to co-operate. People then say: But suppose everybody refused to co-operate, this good would be unobtainable. The refuser will generally be inclined to admit that this is a powerful argument against the rightness of his refusal. We may note further that this line of argument is only employed as a rule when the part contributed to the whole good by each member of the group is very small. We can reinforce this point by a negative instance. Suppose my friend discusses with me the propriety of murdering his wealthy uncle Joseph who has remembered him in his will. Unless I am a Kantian or suspect my friend of being one, I shall not consider it appropriate to point out to him that, if everyone murdered his wealthy uncle from whom he had expectations, a deplorable insecurity would prevail among a deserving class of men and an excuse would be provided for them to leave their money to missionary societies.¹ I should rather insist on the loss of well-being to Uncle Joseph himself, and the probable detection and execution of my friend without the enjoyment of his legacy. And this is because the murder of Uncle Joseph would be in itself an act fraught with appreciably evil consequences, and because for considerable evil to be wrought in this direction the co-operation of many uncle-murderers is not essential.

This fact that the argument is only used where the contribution of each member of the group is necessarily small suggests that the moral microscope plays at any rate some part in these instances. An abstention from joining a group is of course from an ethical point of view just as much an action as joining it. And it may be said that the argument is to show that A's abstention must be wrong though its badness is not obvious on inspection, because a

¹ Though of course a strict Kantian could not use even as suitable an argument as this.

large number of precisely similar abstentions would have admittedly bad results.

But it is surely easy to see that this argument is here a very feeble one. It is quite true that A's abstention *would* have had consequences if it took place together with the abstention of a great many other people. But it does not in the least follow that it *will* have any bad consequences if it take place together with but few other abstentions. Now in judging of the rightness or wrongness of a proposed action it is admitted that we ought in general to take into account the circumstances under which it is to be performed. If a doctor is considering whether he ought to administer chloroform to a patient, it is his duty to consider the particular state of that particular patient's heart, and he will justly be blamed if he omits to do so. Similarly it is impossible to see why A should not have the right and indeed the duty to consider the actual fact that most people have joined the group when he debates on the rightness or wrongness of his joining. It is in fact easy to produce cases where refusal to consider the actual facts about the number of people who have joined will apparently lead A to make a wrong decision. Let us suppose that a group G is co-operating to produce a certain result. Let us suppose that n people have joined the group and let us further suppose that, however great n may be, the joining of the group by an $n+1$ th individuals entails certain sacrifices on him. It is probably reasonable to suppose (a) that the sacrifices made by each individual are lessened as the number of members increases but that the rate of decrease diminishes as n grows greater; (b) that the amount of good produced by the group (apart from the sacrifices) increases as n increases, but that after a certain point the rate of increase diminishes as n grows greater.² If now we call $s(n)$ the

² This hypothesis is unduly favorable to the argument which we are criticising. There are many groups where efficiency would after a certain point continually diminish as n increased. *E.g.*, our drains would be less well cleaned if millions of people were persuaded by the principle of false universalisation to join the group of drain-cleaners. And this would follow from the mere undue increase of numbers in the group, quite apart from the other obvious loss to the general good by this withdrawal from their other occupations.

total sacrifices made where there are n members and $g(n)$ the total good produced by their efforts (apart from the sacrifices) it is quite likely that a point will be reached where

$$g(n+1) - s(n+1) < g(n) - s(n)$$

When this point is reached it would seem to be the duty of people to refuse to join the group, and if they let themselves be guided by the mere fact that $g(0)=0$ and decide to join, they will presumably decide wrongly. Here again the fact that n cannot exactly be determined is merely of practical interest; probably in most cases upper and lower limits could be given for it. We see then that if moral microscopic considerations be the only ones involved in these applications of the principle of false universalisation there is no reason to suppose that the argument to the action of any given person is *ever* valid, and strong reason to believe that after a certain point it will lead to wrong decisions. And in actual fact we notice that the moral microscope is not the main use of the principle in these examples. Suppose that A is told that he ought to co-operate with a certain group because he approves of the end that can be obtained by their efforts, and is reminded that if everybody refused to co-operate the end could not be obtained. If he replies that the end will be obtained whether he co-operates or not and that therefore the hypothesis is as idle as the hypothesis that the moon is made of green cheese, the reply will probably be: But is it fair that other people should do all the work and that you should share in the profits? This seems to be the natural development of the argument from false universalisation in the examples that we are considering where it is clear that the moral microscope is an insufficient explanation. Let us consider it as carefully as possible.

The argument, I think, runs somewhat as follows. You admit that a certain good result can only be obtained by the co-operation of a number of people. Further, this co-operation involves certain sacrifices on the part of all the co-operators. Lastly, the good aimed at is one which,

from the nature of the case, must be enjoyed by all the members of a certain class whether this class be identified with the group of producers or not. The enjoyers may not all be producers and the producers may not all be enjoyers. *E.g.*, if any good results come to the victors in a war they will be of such a kind—national prosperity, feeling of national pride, etc.—that they will *ipso facto* be enjoyed by many members of the victorious nation independently of whether they helped to produce them or not. On the other hand, it is quite certain that many of the producers cannot be enjoyers, because they will be dead or injured for life. A feeling of national pride is *e.g.* a very poor compensation for the loss of both eyes and a leg. Now it may be true that just the same good will be produced whether you co-operate or not, but there is no relevant difference between you and those who join which entitles you to the halfpence without the kicks and them to the certainty of kicks and the possibility of no halfpence.

In fact the argument turns upon distribution. What I mean is this. The possession of certain things and the existence of certain states of mind is intrinsically good. And it is the duty of everyone to aim at what he believes will be the best possible state of affairs on the whole. But the goodness or badness of a complex state of affairs is not a function merely of the goodness or badness of its parts. A certain set of goods distributed in one way between a number of people may constitute an intrinsically better state of affairs than the same set distributed differently. And the appeal to “fairness” seems to rest on the principle that the best possible state of affairs is reached when the group of producers and that of enjoyers is as nearly identical as possible. In fact common-sense would probably go further than this and say that the best possible result was reached when (a) producers and enjoyers are identical and (b) the share in the good produced that falls to each producer is proportional to his sacrifice in producing them. So the argument from fairness really is that the group of producers and the group of enjoyers have no *a priori* or

natural identity; that it is morally desirable that they should be as nearly identical as possible; and that the only way to secure this is for all enjoyers so far as possible to become producers, even though the total product is not thereby increased.

This argument sounds plausible, and I believe that it is in some sense true. But it contains a number of unanalysed difficulties, and it is important to try and determine its limitations. At the very outset a serious difficulty meets us. We have seen that it is almost certain that a point will be reached where, if we disregard questions of distribution, the extra sacrifice due to an additional member joining the group will more than counterbalance the extra benefit due to his action. When we take distribution into account we can see that this limit may very well need to be overstepped in some measure; we can even see that a better total state may be produced by a smaller amount of goods better distributed than by a larger amount worse distributed. But it seems very difficult to deny that there may be a limit beyond which good distribution is too dearly purchased. I imagine that the remark *fiat justitia ruat cælum* would be the denial of this possibility. But I am not sure that anyone would maintain this maxim unless he felt confident that the world is so constituted that the heavens never will fall if justice be done. Now, whether this confidence be justified or not—I myself see no justification for it—it is a metaphysical proposition and not a proposition of pure ethics. We must therefore, I think, be prepared to admit in theory that there may come a point where it is better that some people should refuse to co-operate although this involves an imperfect distribution, than that they should by co-operating produce a much smaller net-balance of goods though perfectly distributed. Once more it is no objection to say that it is totally impossible to determine exactly where this point comes in any particular case. This is quite true, but it is too common a difficulty in ethics to worry us, and we know that we are lucky in ethical questions if we can state upper and

Vol. XXVI.—No. 3.

lower limits that are not too ridiculously far apart. Where right ends and wrong begins between these limits, it is utterly impossible to say.

We may now ask ourselves the question: Is there any necessary connexion between the method of false universalisation and the argument from fairness? We have indeed seen that the former tends to pass into the latter when its more obvious incoherences are pointed out, but *need* we pass to the argument from fairness by way of the argument from false universalisation? I think we may say that, so far as we have yet dealt with the argument from fairness, it has no logical connexion whatever with the argument from false universalization. The only connexion is that both arguments attempt to show that everyone in a certain group should join a certain other group. The argument from false universalisation does this by pointing to the evil consequences of no one joining this group, and it is of course a necessary step in the argument from fairness that it shall be admitted that certain good will result if this group be formed and will not result if it be not formed. But this is the only use that the argument from fairness makes of the argument from false universalisation; it merely takes over one of the premises of that argument. Its own peculiarity is that it supplies two further premises, one factual and one ethical, which the argument from false generalisation does not use. These are (a) the factual premise that the goods are of such a kind that they must be enjoyed by a group not necessarily identical with the group of producers and (b) that the best distribution is one in which the producers and enjoyers of a certain good are identical.

We must now notice a special difficulty which affects the argument from false generalisation. Let A and B be two groups each enjoying a certain good. Let the sole threat to A's good be the existence of a subgroup β in B and the sole threat to B's good be the existence of a subgroup α in A. (*E.g.*, let the good be national prosperity. A and B are two prosperous nations, and the prosperity of A is only

threatened by the armed forces (β) of B, while the prosperity of B is only threatened by the armed forces (α) of A.) A citizen of A refuses to join his army and is told that if everyone did as he proposes to do the prosperity of A will be lost. But this depends on whether he means: I don't intend to join the army of A or I don't intend to join the army of A or of B. If he means the latter, it is clear that the very best results would follow if everyone followed his example, for then both A and B would necessarily retain their prosperity. And if the employer of the argument from false universalisation appeals to the fact that other people certainly will join B's army he is (a) contradicting his own hypothetical premises and (b) laying himself open to the retort that other people will also certainly join A's army. Since it may fairly be assumed that a person who refuses to join his own army intends *ipso facto* to refuse to join any other, his position is absolutely unassailable by the weapon of false universalisation alone. But it is clear that the argument from fairness, if it applies at all, would apply here, too; and therefore we can reinforce our conclusion that there is no essential connexion between the two arguments.

Is there then no valid use for the principle of false universalisation in ethics? I think there is at least one, though it is a very modest one. It can be used to refute a certain kind of mistaken judgment about the rightness of a suggested act. Suppose that certain acts are very unpleasant to everyone and entail very real sacrifices from which everyone shrinks. Suppose further that the performance of such acts by a certain number of persons is essential to the attainment of a considerable good or the avoidance of a considerable evil. If now a man says: I will not act thus *because* I dislike the sacrifice then it is open to us to point out to him that, if this be his sole ground, it is just as valid a ground for all other people, since by hypotheses they all dislike the sacrifice. If then he is right in refusing to do the act, all other people will also be right in refusing on the same ground. But the result will be that a great good will be lost or a great evil suffered.

Now it cannot be the case that the result of a number of right actions can be a state of affairs which can be foreseen to be worse than if people had acted differently. Hence we can conclude that these actions could not all be right. But if his ground for supposing that *his* action was right were valid all these actions would be right. We therefore are forced to conclude, not that his action was wrong (for that does not follow at all), but that at any rate his reasons for supposing it to be right were unsound.

We must now notice the extreme limitations of this use of the principle. (1) It does not prove that *all* the people who abstained acted wrongly; it only proves that some acted wrongly and it offers us no means of deciding which. To decide which one acted wrongly we should have to consider the details of each man and the circumstances under which he made his decision. To decide whether *all* acted wrongly we should have to introduce the principle of fairness, and, as we have seen, it does not necessarily follow from this that all who abstained must have acted wrongly. (2) The argument only holds where the dislike of the sacrifice may be assumed to be nearly equal in everybody. If a man refused to marry because he strongly disliked the idea of marriage you could not prove him to be wrong by the principle of false universalisation; for he might justly say that most other people do not dislike marriage so strongly as he does and therefore his objection could not be universalised. Other people whose dislike was weaker or non-existent could have no ground to conclude that an abstention in them would be right because an abstention in him was right. (3) This brings us to an exceedingly difficult point where even this modest use of the principle threatens to be wrecked. Suppose the dislike of a sacrifice were about equally strong in everyone. We must admit that some people succeed much better than others in overcoming such dislikes from a sense of duty and other motives. Might not a man argue thus? I am not going to do this action because I dislike it and my dislike is stronger than my sense of duty. But this does not

prove that it will be right for others to abstain who have the same dislike but a stronger sense of duty. Hence there is no contradiction between my judgment: This abstention is right in me with my weak sense of duty, and the fact that if all other people whose dislike of the sacrifice is as great as mine were to abstain a great good would be lost or evil produced. For if other people accept my principle it will only justify them in abstaining provided their sense of duty is as weak as mine; and this by hypothesis is not the case. The fact is that you cannot disprove by the method of false universalisation the judgment: My sense of duty is so weak compared with my dislike of the necessary sacrifices that I do right in abstaining from the action.

Indeed, paradoxical as such a judgment seems and greatly as it might be abused, it is not in general refutable. In the first place you certainly cannot say that the actual strength of my desire to do what I believe to be right is *never* a relevant factor in deciding on my proper conduct. For (a) it is certainly clear that the actual strength of *other men's* desire to do what they believe to be right is a very relevant factor in deciding what I ought to do. Most people believe that it is wrong to read private letters, but I know that their desire to do what they believe to be right is very liable to be overcome by their curiosity. And this is a relevant factor in my decision that I ought to keep important private letters under lock and key. But further (b) a man's knowledge that his own desire to keep sober because he believes sobriety to be right is very liable to be overcome by his desire to be drunk when he tastes alcohol is a relevant factor in judging whether he ought or ought not to be a total abstainer. These of course are cases where of two alternatives A and B, A would have better consequences if it were not that it presents a temptation where my sense of duty is weakest, and that the total consequences will be worse if A be performed and the temptation be succumbed to than if B be performed. And it may be admitted that this does not correspond accurately to the situation that led to this discussion.

We may state that situation as follows. A and B are two alternatives open to me, and I believe that A will have better consequences than B. A presents no *especially* dangerous temptation, the only temptation that it presents to my weak sense of duty is the temptation of not completing it. B does not present this temptation so strongly.³ Is the weakness of my sense of duty a relevant factor here? It seems to me that it very well may be. It is no doubt true to say that we are here turning our attention from the probable consequences of our actions to the probability of our performing an action. But, since the consequences will not follow unless we do the action, the probability of the consequences must depend on the probability that we shall carry the action through, and this depends on the strength of my sense of duty as against the particular sacrifices that the action involves. If the consequences of A be better than those of B and those of B better than those of the absence or partial completion of A, and if the weakness of my own sense of duty in the face of the special sacrifices involved in A make it more likely that I shall complete B than A, it is my duty to choose B rather than A; and the relevant factor is the weakness of my sense of duty. We see then how very limited is that principle which some men seem to have thought the basis of ethics.

There remains, however, one further question to discuss. We have all along assumed that the rightness or wrongness of an action depends wholly on its actual or probable consequences. Now this seems to me not to be true. It is quite certain that its rightness is a function of the goodness of its probable consequences, but, as I have tried to prove in an earlier paper in this JOURNAL,⁴ its rightness is not entirely determined by this. The rightness seems to be a function of the intrinsic goodness of its motive and of the goodness of its probable consequences. Moreover, it is

³ An example would be the case of a man who was considering whether he ought to enlist in view of the fact that his sense of duty might not be strong enough to prevent him from running away in battle.

⁴ The Doctrine of Consequences in Ethics, xxiv, 293.

certain that in many judgments when we appear to be judging about the rightness or wrongness of an act we are really judging simply about the goodness or badness of a motive. It is therefore necessary for the sake of completeness to ask whether the principle of false universalisation is of any use as a means of judging about motives. It might conceivably have either of two uses. It might (a) help us to recognize what our motives really are, which is of course an essential preliminary to passing any reasonable judgment of value upon them. Or (b) it might help us to judge whether our motives once recognised are good or bad. Moreover, while it is certain that motives have value as means, it seems to me possible that they have intrinsic value, and tolerably certain that they can alter the intrinsic value of wholes consisting of themselves and their consequences. So that (b) divides into three questions. Does the principle help us to judge the goodness of a motive (i) as a means, (ii) as an end, or does it enable us to see (iii) whether its presence adds to or lessens the intrinsic value of a whole composed of itself and a certain set of consequences?

(a) If I propose to perform a certain act and then am asked to contemplate the hypothesis that everyone acts as I propose to do, it will be necessary for me to be quite clear how I *do* propose to act and why. I shall have every reason to try and be clear on this point because otherwise the universalisation will not be fair. Hence if I am challenged to make the universalisation, and especially if I object to the way in which my interlocutor makes it for me, I shall have a very good chance of recognising what my motive really is.

(b) i. The principle of false generalisation may very well show me that my motive is (in a certain special sense which we will explain in a moment) bad as a means. It will not indeed prove to me that my motive in my circumstances leads to worse results than a different motive; but it may show me that it is bad as a **means** in the sense that it would be very unfortunate if it were at all a common

motive with people. Now this does not prove that it actually leads to bad results in *any* case, for you cannot argue from the fact that the *coexistence* of this motive in a great many people would lead to bad results that its existence in any particular person will lead to bad results. Thus it is only in a very special and not very useful sense that the principle will show that a motive is bad as a means. It is very easy to commit a gross fallacy here. This is to confuse the two statements "the motive *m* frequently leads to bad results" and "the frequency of the motive *m* would lead to bad results." From the former we could conclude that any particular instance of the motive is likely to lead to bad results; from the latter we can conclude nothing of the sort. And it is only the latter that could be proved by the principle of false universalisation.

(b) ii and iii. I think that the principle has very little direct connexion with the intrinsic goodness of motives or with the goodness which their presence adds to or subtracts from that of other wholes. But there is a rather roundabout connexion of the following kind. There happens to be one motive—the desire to do what is believed to be right and reasonable as such—which is (a) intrinsically good and (b) adds to the goodness of a great many (though not I think of all) sets of consequences produced by it. Now it is on the whole true of the world as at present constituted that this motive also has beneficial results in the great majority of cases, and that the results of everybody acting from it would be very beneficial. (This would of course be false if the world were so constituted that the more people tried to judge dispassionately about the goodness of the results of certain actions the more they were likely to be mistaken. But there seems no reason to take such a depressing view as this.) Hence if it can be shown that the results of a number of people acting as I propose to act from my motives would be worse than if they acted differently there is at least a strong presumption that my motive is not the desire to do what *is* right and reasonable. Of course even if my motive were a different one it need

not be a bad one. But at least the argument supplies some presumption that my motive does not belong to the most important class of motives which are recognized to be intrinsically good and to add greatly to the goodness of the wholes that are connected with them as consequences.

Such then are the modest and doubtful functions of the principle of false universalisation. Most of its alleged uses we have seen are sheerly fallacious; and often where we cannot prove a fallacy we can see that there is very likely to be one and can produce no clear case where it is quite certain that no fallacy lurks. The most important use is to lead us to the principle of fairness, and yet there seems to be very little logical connexion between the two; so that it is rather a matter of psychology than of logical necessity that the principle of fairness is generally introduced by that of false universalisation. And the principle of fairness itself bristles with difficulties which I have barely touched, but to which I hope to return in a later paper.

C. D. BROAD.

UNIVERSITY OF ST. ANDREWS, SCOTLAND.