Oxford University Press, Amen House, London E.C. 4
GLASGOW NEW YORK TORONTO MELBOURNE WELLINGTON
BOMBAY CALCUTTA MADRAS KARACHI CAPE TOWN IBADAN
Geoffrey Cumberlege, Publisher to the University

LOGIC AND THE BASIS OF ETHICS

BY

ARTHUR N. PRIOR
Lecturer in Philosophy
Canterbury University College, New Zealand

OXFORD
AT THE CLARENDON PRESS

For which of you, intending to build a tower, sitteth not down first, and counteth the cost, whether he have sufficient to finish it?

LUKE 14: 28

FIRST EDITION 1949

REPRINTED LITHOGRAPHICALLY IN GREAT BRITAIN
AT THE UNIVERSITY PRESS, OXFORD
FROM SHEETS OF THE FIRST EDITION
1956

CONTENTS

I

THE NATURALISTIC FALLACY: THE LOGIC OF ITS REFUTATION

If there is any contribution to moral philosophy which is more likely than any other to become permanently associated with the name of Professor G. E. Moore, it is the identification and refutation, in his *Principia Ethica*,[1] of what he calls the 'naturalistic fallacy'. I propose now to explain what it is to which Professor Moore gives this name, and what he considers to be involved in its fallaciousness; and I shall offer reasons for regarding his argument, not as disproving ethical naturalism itself, but as exposing an inconsistency into which some naturalists have fallen.

What Professor Moore means by the 'naturalistic fallacy' is the assumption that because some quality or combination of qualities invariably and necessarily accompanies the quality of goodness, or is invariably and necessarily accompanied by it, or both, this quality or combination of qualities is *identical* with goodness. If, for example, it is believed that whatever is pleasant is and must be good, or that whatever is good is and must be pleasant, or both, it is committing the naturalistic fallacy to infer from this that goodness and pleasantness are one and the same quality. The naturalistic fallacy is the assumption that because the words 'good' and, say, 'pleasant' necessarily describe the same objects, they must attribute the same quality to them. We might, with Mill, call the objects to which a term is applicable the denotation of the term, and the characteristics which an object must have for the term to be applicable to it, the connotation of the term.[2] What the man who commits the naturalistic fallacy

[1] pp. 6–17.

[2] J. S. Mill, *System of Logic*, I. ii. 5. The importance of Mill's distinction in the interpretation of Professor Moore's account of the

fails to realize is that 'good' and some other adjective may denote or be applicable to the same things, and yet not connote the same quality, i.e. describe the things in the same way. The difference between identity of denotation and identity of connotation may be brought out, as Professor Moore shows, by the following simple consideration: If the word 'good' and, say, the word 'pleasant' apply to the same things, but do not attribute the same quality to them, then to say that what is pleasant is good, or that what is good is pleasant, is to make a significant statement, however obvious its truth may appear to many people. But if the word 'good' and the word 'pleasant' not merely have the same application but the same connotation or 'meaning'—if, that is to say, the quality of pleasantness is identical with the quality of goodness—then to say that what is good is pleasant, or that what is pleasant is good, is to utter an empty tautology, or, as Mill would call it,[1] a 'merely verbal' proposition; for both statements are on this supposition merely ways of saying that what is pleasant is pleasant.

From this consideration Professor Moore attempts to show that the term 'good' is incapable of definition. By 'definition' he means the exhibition of a quality referred to by some term as a combination of simpler qualities. And he argues that if we take any such combination of relatively simple qualities (such as the combination 'being what we desire to desire'), the statement that what possesses this combination of qualities is good (e.g. the statement that what we desire to desire is good) will always be found on careful inspection to be a significant statement and not a mere truism (like 'What we desire to desire, we desire to desire'). But this is not all that he claims to be able to show by this method. We may use

naturalistic fallacy is rightly emphasized in Dr. D. Daiches Raphael's *The Moral Sense*, pp. 111–14; though on p. 113 Dr. Raphael attributes to certain modern mathematicians a confusion in regard to this point, of which I do not think they are really guilty.

[1] *System of Logic*, I. vi.

it, he thinks, to show that goodness is not only simple, i.e. incapable of analysis into simpler parts, but unique. For even if we take a *simple* quality, such as pleasantness, we can always see that it is significant, and not a mere truism, to assert that what possesses this quality is good. (Despite his definition of 'definition', as analysis, he slips readily into calling 'Good means pleasant' a 'definition' too.)

This latter contention of Professor Moore's is exceedingly difficult to state with any precision. It plainly does not apply to the quality of goodness itself—it *is* a truism to assert that what is good is good. Nor does it apply to the quality of goodness itself when it is merely given another name, such as 'value' (which is often used as synonymous with 'goodness' by Professor Moore, as well as by many other writers). Yet if we merely say that goodness is not identical with any other quality, this is itself a truism—it merely tells us that goodness is not identical with any quality, simple or complex, with which it is not identical. It is a little ominous that Professor Moore quotes on his title-page the sentence from Bishop Butler, 'Everything is what it is, and not another thing'. For who would deny this? Even the man who identifies goodness with pleasantness, i.e. who regards 'good' as a mere synonym of 'pleasant', would not deny that it is in this sense 'unique'. For pleasantness also 'is what it is, and not another thing'; and to say that goodness is pleasantness is not, on such a view, to deny that it is what it is, or to affirm that it is another thing—it is merely to deny that pleasantness is 'another thing'.

Professor Moore's real aim, of course, is to show that goodness is not identical with any 'natural' quality. This is why he calls the kind of identification which he is opposing the 'naturalistic' fallacy. But what does he mean by a 'natural' quality? He attempts an answer to this in the *Principia*, but now says that the answer there given is 'utterly silly and preposterous',[1] as indeed it is (there is no need to reproduce

[1] *The Philosophy of G. E. Moore*, p. 582.

it here). And at times it looks very much as if what he means by a 'natural' quality is simply any quality other than goodness or badness, or at all events other than goodness, badness, rightness, wrongness, and obligatoriness (if the last three are taken to be distinct from goodness and badness and from one another—which, in the *Principia Ethica*, they are not), and compounds containing these. But if this is what he means, are we not back where we were?—are we not still left with the truism that 'Everything is what it is, and not another thing'?

It is worth examining this sentence in its original context. Butler's argument in the paragraph from which it is taken[1] is directed against people who were putting it about that it can never be to any man's interest to be virtuous, since disinterestedness is of the essence of virtue. Mandeville, holding that nothing is virtuous but 'self-denial', went so far as to say that virtue was not only not in a man's own interests but generally not in anyone else's either, so that 'private vices' were 'public benefits'. So Butler sets out to show that virtue and disinterestedness are not the same thing (though virtue and self-interest are not the same thing either).

'Virtue and interest, are not to be opposed, but only to be distinguished from each other; in the same way as virtue and any other particular affection, love of arts, suppose, are to be distinguished. Everything is what it is, and not another thing. The goodness or badness of actions does not arise from hence, that the epithet, interested or disinterested, may be applied to them, any more than that any other indifferent epithet, suppose inquisitive or jealous, may or may not be applied to them; not from their being attended with present or future pleasure or pain; but from their being what they are; namely, what becomes such creatures as we are, what the state of the case requires, or the contrary.'

Butler is not, I think, denying that the moral quality of an

[1] *Sermons on Human Nature*, Preface, par. 39.

act is determined by its other qualities—he is not denying, for example, that in a given situation a certain intensity of jealousy is always wrong, i.e. 'unbecoming' to 'such creatures as we are'. But he is denying that anything of this sort—expressing jealousy of such-and-such an intensity in such-and-such a situation—is what we *mean* by calling an act good or bad. Its goodness or badness is its 'moral appropriateness' to our nature and our situation. That is, its goodness or badness is its goodness or badness; it is its 'being what it is', good or bad as the case may be. Goodness or badness cannot be identified with any 'indifferent' epithets.

But what kind of epithet is that? If we take 'indifferent' to mean merely 'non-moral'—i.e. if an 'indifferent' epithet is any one that does not mean the same as 'good' or 'bad'—is not Butler's argument open to the same objection as Professor Moore's? Certainly goodness and badness are not to be identified with any qualities that are other than goodness and badness; but how does this forbid us to identify goodness with disinterestedness? Does not the identification of goodness with disinterestedness merely remove the latter from the class of 'indifferent' epithets, i.e. from the class of the 'other things' which goodness is not (just as, on Butler's own view, what we have called 'moral appropriateness' is something that goodness is, 'and not another thing')?

I think we must take it that what Butler means by 'indifference', and Professor Moore by 'naturalness', is something more than mere non-identity with goodness or badness. Their view seems to be that all qualities other than goodness and badness have something positive in common—something which is so near to universal that we do not notice it until we compare the qualities marked by it with goodness and badness; and then it is intuitively evident. When we compare such qualities as goodness and badness with such qualities as pleasantness, pinkness, everlastingness—to take a quite random selection—we see that the former and the latter are

not only individually non-identical, as pleasantness and pinkness are, but fall into two quite different categories or 'realms', namely, those which we sometimes call the realm of value (or of duty) and the realm of fact. These terms are not perhaps quite fortunately chosen, since it may be held—it is held by Professor Moore, for example, and was by Butler—that to say that something is our duty, or possesses value, is to state a fact, albeit of a very peculiar kind. (We shall see in the sixth and seventh studies that there are some writers who deny this; but such a denial seems to amount to saying that there is really only one realm—the 'natural' one—and this is not the position which we are at present trying to formulate.) But however we describe these two 'realms', their existence and distinctness is what seems to be referred to in Professor Moore's distinction between ethical predicates and all 'natural' ones, as it is in the old distinction betweeen the 'moral' perfections of the Deity and His 'natural' ones (omnipotence, omniscience, eternity, &c.), and in Aristotle's distinction between the 'ethical', the 'natural', and the 'logical' fields of inquiry. And Aristotle notes that 'the nature of each of the aforesaid kinds of proposition is not easily rendered in a definition, but we have to try to recognize each of them by means of the familiarity attained through induction, examining them in the light' of certain 'illustrations' given previously—'ethical' questions being illustrated by 'Ought one rather to obey one's parents or the laws, if they disagree?' and 'natural' ones by 'Is the universe eternal or not?'[1] (Aristotle is here using 'induction' to mean, not a process of reasoning, but the examining of instances until their common quality 'dawns' upon one—his appeal is to intuition.) But such an intuitively perceived difference between 'moral' qualities and all others plainly goes far beyond anything that can be proved from the principle that 'Everything is what it is and not another thing', since this principle would

[1] *Topics*, 105^{b}21–9.

still apply within a single 'natural' realm even if there were no other.[1]

Professor Moore's appeal to this truism, and the little dialectical device which he bases upon it, are not, however, entirely pointless. For there *are* occasions when men implicitly deny logical truisms, and need to be reminded of them; namely, when they are inconsistent. It is not against the naturalist as such, but the inconsistent naturalist, the man who tries to 'have it both ways', that Professor Moore's type of argument is really effective and important. And such people are not uncommon. Professor Moore himself mentions them—the people who begin by laying it down as a truth of primary importance, perhaps even as something rather revolutionary, that nothing is good but pleasure, or that nothing is good but what promotes biological survival, and who, when asked why they are so certain of this, reply that 'that is the very meaning of the word'. To such people it is certainly legitimate and necessary to reply that if pleasantness, or the promotion of survival, is what 'goodness' *means*, then the fact that only pleasure is good, or that only what promotes survival is good, is hardly worth shouting from the house-tops, since nobody in his senses ever denied that what is pleasant, and only what is pleasant, is pleasant, or that what promotes survival, and only what promotes survival, promotes survival. What these people would plainly like to hold is that goodness is both identical with pleasantness and not identical with it; and, of course, it cannot be done. They want to regard 'What is pleasant is good' as a significant assertion; and it can only be so if the pleasantness of what is pleasant is one thing, and its goodness another. On the other hand they want to make it logically impossible to contradict this assertion—they want to treat the opposing assertion that what is pleasant may not be good as not merely false but logically

[1] This point is elaborated in an article on 'The Naturalistic Fallacy', by W. K. Frankena, in *Mind*, 1939, pp. 472 ff.

absurd—and this can only be done if pleasantness and goodness are taken to be identical. To represent an opponent's position in such a way as to make it not only false but self-contradictory is a dialectical triumph which can never be obtained without being duly paid for; and the price is the representation of one's position as not only true but a truism. 'If a denial is to have any value as a statement of matter of fact', as Dr. J. N. Keynes says,[1] then what it denies 'must be consistent with the meaning of the terms employed. . . . The denial of a contradiction in terms . . . yields merely what is tautologous and practically useless.'

It is sometimes pointed out by naturalists that there is never more than one ethical statement which is rendered trivial by a naturalistic definition of 'good'. If, for example, we use 'good' as synonymous with 'conducive to biological survival', then, while it is a truism to say that what is conducive to biological survival is 'good' in this sense, it is not a truism to say that pleasure is, since it is not a truism to say that pleasure is conducive to survival. We shall find shortly that there is a point at which this consideration is important; but if Professor Moore's argument is regarded as a criticism of the attempt to deduce significant assertions from definitions, this answer to it is irrelevant, since the statement which the definition makes trivial is always precisely the one which it is put forward to 'prove', in a sense in which it is not trivial but significant. A man who has defined 'good' as 'conducive to biological survival', with the express purpose of establishing it as an ethical principle of primary importance that only what conduces to survival is good, will not be greatly cheered by the consideration that it is 'only' this principle which the definition renders insignificant.

Confronted with Professor Moore's argument, an inconsistent ethical naturalist has two courses open to him. He may clear himself of inconsistency, on the one hand, by abandoning his naturalism—he may continue to insist that only pleasure, or conduciveness to survival, or whatever it may be, is good, but may preserve the significance of this assertion by sacrificing its certainty, admitting that its denial, though still in his opinion false, is not self-contradictory. Professor Moore writes as if this is what any naturalist who really grasps his argument will do—he seems to consider his argument a *refutation* of naturalism. But a naturalist can preserve his naturalism if he wants to, even in the face of Professor Moore's argument—he can do so by admitting that the assertion that, say, pleasure and nothing but pleasure is good, *is* for him a mere truism; and that if Ethics be the attempt to determine what is in fact good, then the statement that what is pleasant is good is not, strictly speaking, an ethical statement, but only a way of indicating just what study is to go under the name of 'Ethics'—the study of what is actually pleasant, without any pretence of maintaining that pleasure has any 'goodness' beyond its pleasantness. He might add at the same time that he is not only not going to discuss goodness as a 'non-natural' quality, but that in his belief there is no such quality, and that this is worth shouting from the housetops, as it liberates us from a transcendental notion which has haunted us too long. (He might say that this is what he really means by the assertion that 'Nothing is good but pleasure'—he means, not that what is pleasant alone possesses some other quality called 'goodness', but that there are no qualities beyond 'natural' ones such as pleasantness to which the word 'goodness' could be applied.) Indeed, he is bound to say something of this sort if he is to justify his appropriation of the word 'good' for the purpose to which he puts it. And such a man, it seems to me, should be prepared to state his position in an alternative way, namely, as a denial that there *is* such a study as Ethics—he should be prepared, for the sake of clarity, and to further the mental 'liberation' in which he is primarily interested, to call his inquiry into the

[1] *Formal Logic*, pp. 119–20.

sources of pleasure, not Ethics, but some such name as 'Hedonics'; or if he defines goodness as 'conduciveness to survival', to call his substitute for Ethics 'Biological Strategy'.

But how—as Mr. E. F. Carritt pertinently asks[1]—can we be 'liberated' from a notion which we cannot ever have had? For how can we have had a 'transcendental' notion of goodness if the word which is alleged to have called it up is also alleged to have no meaning, or none beyond ones which are not 'transcendental' at all? Even this question it is not beyond the power of a consistent naturalist to answer.

'A name', as J. S. Mill points out, 'is not imposed at once and by previous purpose upon a *class* of objects, but is first applied to one thing, and then extended by a series of transitions to another. By this process . . . a name not unfrequently passes by successive links of resemblance from one object to another, until it becomes applied to things having nothing in common with the first things to which the name was given; which, however, do not, for that reason, drop the name; so that it at last denotes a confused huddle of objects, having nothing whatever in common; and connotes nothing, not even a vague and general resemblance. When a name has fallen into this state, . . . it has become unfit for the purposes either of thought or of the communication of thought; and can only be made serviceable by stripping it of some part of its multifarious denotation, and confining it to objects possessed of some attributes in common, which it may be made to connote.'[2]

And this, a naturalist may say, is precisely what has happened with the word 'good', and what needs to be done about it. At present, when we call a thing good we may mean that it is pleasant, or that it is commanded by someone, or that it is customary, or that it promotes survival, or any one of a number of things; and because we use the same term to connote all these characteristics, we think there must be some other single characteristic which they all entail; but in

[1] *Ethical and Political Thinking*, pp. 33–4.
[2] *System of Logic*, I. viii. 7; see also IV. iv. 5, v. 2.

fact there is not. When it is said that being good means promoting survival, we are dissatisfied; we feel that it is still significant to say that promoting survival is good; and the same thing happens with every identification that is suggested; but this is just because, in each case, the other meanings are still hovering in our minds—to say that promoting survival is good is significant because it means that to promote survival is what we desire; to say that what we desire is good is significant because it means that what we desire promotes survival; and so on. Once we realize this, we may either recommend and adopt a more consistent usage; or we may leave the word with its present 'flexibility', but with the misleading suggestions of that flexibility removed. The naturalist who proposes some unambiguous definition is taking the first course.[1]

This way of dealing with words like 'good' is characteristic of the 'therapeutic positivism' developed at Cambridge in the past few decades under the influence of Professor Wittgenstein. While this is unquestionably a useful philosophical technique, there are obvious limits to its applicability. For it is plain that in some cases in which diverse objects are called by a common name there *is* a common characteristic on account of which the name is given to them all. We need some principle enabling us to decide when such a common characteristic exists and when it does not; and what principle we use for this purpose will depend upon our general philosophical position. Analyses of the sort just given cannot therefore replace philosophical inquiry, as 'therapeutic positivists' seem at times to think they can, but both aid it and depend upon it. If we have other reasons for regarding the distinction between the 'natural' and the 'moral' realms as an illusory one, then tricks of language may explain how the

[1] For an answer to Professor Moore along these general lines see a dialogue by E. and M. Clark entitled 'What is Goodness?' in the *Australasian Journal of Psychology and Philosophy*, 1941.

illusion has come about; but it may still, as a matter of fact, be real.

It remains true, however, that a naturalist *can* extricate himself from Professor Moore's trap if he is bold enough and tough enough. And in imagining that in his refutation of what he calls the 'naturalistic fallacy' he has refuted naturalism, Professor Moore has himself fallen into a fallacy not unlike it. For if Professor Moore's own non-naturalism is a significant belief, then it must be possible to formulate the naturalism which it contradicts in a significant way; and if naturalism itself, and not merely the inadvertent combination of naturalism with something inconsistent with it, is senseless, then the denial of it is trivial. A significant non-naturalism, in other words, must comprise more than mere freedom from the 'naturalistic fallacy'.

II

THE AUTONOMY OF ETHICS: (1) CUDWORTH

THE same broad type of moral philosophy as Professor Moore has taught at Cambridge in our own time was also taught there in the seventeenth century by that difficult but rewarding writer Ralph Cudworth. Cudworth is mentioned by Rashdall[1] as anticipating Professor Moore's opinion that 'good is indefinable'; and, as we shall see shortly, he defended this opinion by the same bad argument. But like Professor Moore he also had a good argument against ethical naturalists who could not be quite consistent; and though it was not quite the same argument as Professor Moore's, it is equally worthy of our attention, and we shall accordingly study it, indicating in later studies how it was developed by later and clearer writers.

The inconsistent ethical naturalism which Cudworth criticized took the form of an identification of goodness or rightness (as I have already indicated, it is not necessary for my present purpose to distinguish sharply between these) with obedience to someone's will—the civil sovereign's or God's—coupled with an insistence, as if it were an insistence on something of the first importance, that to obey this person is good or right, and to disobey him bad or wrong—an insistence, in short, that we have in some significant sense a duty to obey him. Hobbes, in particular, sometimes spoke in this way about the civil ruler, and Descartes and various theologians about God. Against all attempts to make goodness thus completely dependent on a superior's will, Cudworth argues that

> 'Moral good and evil . . . cannot possibly be arbitrary things, made by will without nature; because it is universally true, that things are what they are, not by will but by nature. As for ex-

[1] *The Theory of Good and Evil*, vol. i, p. 136.

the facts, about which it is supposed, rightly or wrongly, that a majority of careful observers are not likely to be mistaken.

We may approach the matter in another way. It is a very important part of Professor Findlay's technique to define ethical responses as ones which we are prepared to have tested in these ways—he says that we simply would not *call* a response an ethical one if the person making it were not prepared to submit it to such and such a test. And we must admit that this procedure, like all 'naturalistic' definition of key ethical concepts, is legitimate, in so far as men, and also societies and cultures and schools of thought, may use words as they please. But, as Professor Moore insists, whoever lays down a definition must be prepared to accept its consequences. Now Professor Findlay tells us at one point that 'the moral sphere is really one of these spheres in which the *orbis terrarum* may be said to judge securely'.[1] But since he defines an ethical response as one in which we are prepared to submit to the decision of the *orbis terrarum*, this is simply a tautology. And I do not think Professor Findlay would attempt to deny that it is; but one further consequence of its being a tautology which he appears to have overlooked is that it makes his employment of this particular test for the 'truth' of an ethical response differ *in toto* from its employment as a test of the truth of a judgement. For if it is a fact that the *orbis terrarum* judges securely on questions of truth and falsehood (as those terms are ordinarily understood), then it is a highly significant fact, and not the mere tautology that the *orbis terrarum* judges as it does. And conversely, if it is a mere tautology that the *orbis terrarum* judges securely in moral matters, then this 'judgement' is not one of truth and falsehood, in any ordinary sense of those terms.

[1] p. 160.

IX

THE NATURALISTIC FALLACY: THE HISTORY OF ITS REFUTATION

We have seen that the claim to infer significant ethical propositions from definitions of ethical terms, which appears to constitute the essence of what Professor Moore calls the naturalistic fallacy, is a special case of a more general fallacious claim, namely, the claim to deduce ethical propositions from ones which are admitted to be non-ethical. We have considered some of the forms in which this claim has been historically put forward, and some of the ways in which it has been historically refuted. We have also considered attempts to give ethics a 'foundation' by misleading extensions of the concept of 'truth', and the ways in which the fallacies involved in such attempts have been or may be exposed. All this has provided us with a broad context in which we can study the history of the exposure, by the method which we now think of as Professor Moore's, of the naturalistic fallacy itself.

The closest approach to an anticipation of Professor Moore that we have yet encountered is perhaps Cudworth's relegation to a parenthesis, as something which his opponents cannot have seriously meant to maintain, of the view that good and evil are 'mere names without signification, or names for nothing else but willed and commanded'. But Cudworth does not explain why he considers this possibility out of the question. Here and there, however, among those who came after Cudworth, there are to be found writers who do consider it worth while to explain why this possibility cannot be seriously entertained. The earliest of such explanations which I have been able to trace is that of Shaftesbury, who points out that 'whoever thinks there is a God, and pretends

formally to believe that he is just and good, must suppose that there is independently such a thing as justice and injustice, truth and falsehood, right and wrong, according to which he pronounces that God is just, righteous, and true. If the mere will, decree, or law of God be said absolutely to constitute right or wrong, then are these latter words'—i.e. the 'pronouncement' that God is just, righteous, and true—'of no significancy at all'.[1] And the anticipation of Moore is made complete a little later by Hutcheson, who writes: 'To call the laws of the Supreme Deity good, or holy, or just, if all goodness, holiness and justice be constituted by laws, or by the will of a superior any way revealed, must be an insignificant tautology, amounting to no more than this, "That God wills what he wills".'[2]

The inconsistent ethical naturalist whom Shaftesbury and Hutcheson had chiefly in mind when formulating their criticism was probably John Locke.

'Things are good and evil', Locke held, 'only in reference to pleasure and pain. That we call good, which is apt to cause or increase pleasure, or diminish pain in us; or else to procure or preserve us the possession of any other good or absence of any evil. And, on the contrary, we name that evil which is apt to produce or increase any pain, or diminish any pleasure in us; or else to procure us any evil, or deprive us of any good.'[3]

And again, 'Good and evil . . . are nothing but pleasure and pain, or that which occasions or procures pleasure or pain to us.'[4] 'Moral' good or evil is a special kind of source of pleasure and pain, namely, 'the conformity or disagreement of our voluntary actions to some law, whereby good or evil'—i.e. pleasure or pain—'is drawn on us by the will and power

[1] *An Inquiry concerning Virtue*, Bk. I, Part III, sect. ii.

[2] *An Inquiry concerning the Original of our Ideas of Virtue and Moral Good*, VII. v; Selby-Bigge, 173. I have drawn attention to this passage, and to the one from Shaftesbury, in the *Australasian Journal of Psychology and Philosophy*, December 1946, p. 172.

[3] *Essay concerning Human Understanding*, II. xx. 2. [4] II. xxviii. 5.

of the law-maker'.[1] And the law which determines what actions are 'sins or duties' (and not merely 'criminal or innocent' in the eyes of one's government, or proper or improper in the eyes of one's society) is the law or commandment of God.[2] Here is 'naturalism' at its purest—'moral' good and evil (right and wrong) reduced to a form or function of 'natural' good and evil (pleasure and pain) in so many words. And Locke is consistent enough to recognize that this makes some propositions about our duty tautologous, or almost so.

'If virtue be taken for actions conformable to God's will, or to the rule prescribed by God, which is the true and only measure of virtue when virtue is used to signify what is in its own nature right and good: then this proposition, "That virtue is the best worship of God", will be most true and certain, but of very little use in human life, since it will amount to no more but this, viz. "That God is pleased with the doing of what he commands"; which a man may certainly know to be true, without knowing what it is that God doth command, and so be as far from any rule or principle of his actions as he was before.'[3]

This is directed against Lord Herbert of Cherbury's claim that the proposition in question is a self-evident ethical principle; Locke in effect admits its self-evidence, but denies its status as an ethical principle. Yet Locke can also say that God has a 'right' to rule us, and that not only because He has power to enforce His commands 'by rewards and punishments of infinite weight and duration', but also because 'he has goodness and wisdom to direct our actions to that which is best'.[4] Does God's 'goodness' here mean merely that He is a source of pleasure? If so, He is also 'evil', as being, to the disobedient, a source of pain. Or is His goodness 'moral'? Then it means, as Hutcheson says, no more than that His will accords with itself. And does His 'right' to impose laws mean that it is in accordance with His laws that He should

[1] Ibid. [2] II. xxviii. 7. [3] I. iii. 18. [4] II. xxviii. 8.

impose them? This also seems to tell us no more than that He commands what He commands.

In the middle of the eighteenth century this criticism was explicitly directed against Locke by Richard Price. '*Mr. Locke* . . . represents *rectitude* as signifying conformity of actions to some rules or laws; which rules or laws, he says, are either *the will of God*, the decrees of the magistrate, or *the fashion of the country*: From whence it follows, that it is an absurdity to apply *rectitude* to rules and laws themselves' or 'to suppose the *divine* will to be directed by it.' 'But', he adds, 'it is undoubted that this great man would have detested these consequences; and, indeed, it is sufficiently evident, that he was strangely embarrassed in his notions on this, as well as some other subjects.'[1] Price makes a similar criticism of Bishop Warburton, who 'maintains, that moral obligation always denotes some object of will and law, and implies some obliger. Were this true, it would be mere jargon to mention our being obliged to obey the Divine will; and yet, this is as proper language as any we can use.'[2] These are, of course, mere hints of Professor Moore's argument from trivialization, like the hint in Shaftesbury; not full anticipations of it, like that in Hutcheson. But the latter may be found in Price too; in fact, no other writer has anticipated Professor Moore quite so completely.[3]

This more complete anticipation occurs, curiously enough, in a section[4] in which Price's main purpose is to state his difference from Hutcheson; though it occurs there as a digression. Their difference concerns what Price calls 'the Foundation of Morals'. On Hutcheson's view (which Price identifies perhaps too unreservedly with Hume's), 'moral right and wrong, signify nothing *in the objects* themselves to which they

[1] *Review*, p. 43; Selby-Bigge, 609.
[2] *Review*, p. 116; Selby-Bigge, 684.
[3] The first person to have noticed this, so far as I am aware, was Dr. Raphael. See *The Moral Sense*, pp. 1, 111 ff.
[4] *Review*, I. i.

are applied, any more than agreeable and harsh', and 'our perception of *right*, or moral good, in actions, is that agreeable *emotion*, or feeling, which certain actions produce in us: and of wrong, the contrary'.[1]

'The present enquiry therefore is; whether this be a true account of virtue or not; whether it *has* or has *not* a foundation in the *nature* of its object; whether *right* and *wrong* are real characters of *actions*, or only qualities of our *minds*; whether, in short, they denote what actions *are*, or only *sensations* derived from the particular frame and structure of our natures.'

But there is one set of theories—'the schemes which found morality on self-love, on positive laws and compacts, or the Divine will'—which may not seem to fit very well into either of these pigeon-holes. But these 'must either mean, that moral good and evil are only other words for *advantageous* and *disadvantageous*, *willed* and *forbidden*. Or they relate to a different question; that is, not to the question, what is the nature and true *account* of virtue; but what is the *subject-matter* of it.'[2] If the latter is their meaning—if they mean that being advantageous to the agent, or being commanded by God or by some other authority, are the only characteristics of actions which *make* them right—then what they have to say has no bearing on the question as to what broad kind or category of quality 'rightness' itself is. On the other hand, if the key propositions of these schemes are intended as definitions, the consequence of accepting them would be that 'it would be palpably absurd in any case to ask, whether it be right to obey a command, or wrong to disobey it; and the propositions, obeying a command is right, or producing happiness is right, would be most trifling, as expressing no more than that obeying a command, is obeying a command, or producing happiness, is producing happiness'.[3] Here we have Professor Moore's whole armoury—not only

[1] Selby-Bigge, 585.
[2] Ibid., 586.
[3] Ibid., 587.

the argument from trivialization, but the distinction underlying it, between a definition of a moral term, and a significant ethical generalization.

Views akin to Locke's found a number of adherents among the eighteenth-century clergy. Fielding's *Tom Jones* contains a number of debates between one such clergyman, who is given the name of 'Thwackum', and a person called 'Square', who is of the school of Clarke (Fielding himself fairly plainly writes from the point of view of Shaftesbury—a moderate 'sentimentalism', more concerned to insist upon the reality of generous emotions in men than to propound any theory of the nature of the moral faculty). In 1785 this 'theological naturalism' was given its classical form in Paley's *Principles of Moral and Political Philosophy*. Paley defines virtue as 'the doing good to mankind, in obedience to the will of God, and for the sake of everlasting happiness';[1] 'right' as 'consistency with the will of God';[2] and 'obligation' as being 'urged by a violent motive resulting from the command of another'.[3] If it is asked, 'Why am I *obliged* to keep my word?' on this system, 'the answer will be, "because I am urged to do so by a violent motive" (namely the expectation of being after this life rewarded, if I do, or punished for it if I do not), "resulting from the command of another" (namely of God)'. 'This solution', Paley complacently comments, 'goes to the bottom of the subject, as no further question can reasonably be asked. . . . Private happiness is our motive, and the will of God our rule.'[4]

In a note[5] to the final edition of his *Review*, Price makes the comment on this that might have been expected from him. 'Mr. PALEY'S definition of RIGHT is, "the being consistent with the will of God". RECTITUDE, therefore, can be no guide to God's will itself; and to say that his will is a righteous will, is the same with saying that his will is his will.' The same consequence of Paley's position is emphasized in the critical

[1] I. vii. [2] II. ix. [3] II. ii. [4] II. iii. [5] Note F.

annotations to the edition of the *Moral and Political Philosophy* which was produced in the middle of last century by Archbishop Whately, the logician. Whately notes, to begin with, that if a man 'attaches no meaning to the words "good", and "just", and "right", except that such is the divine command, then, to say that God is good, and his commands just, is only saying in a circuitous way, that He is what He is, and that what He wills He wills, which might be said of any Being in the universe'.[1] Whately, as a theist, admits that

> 'we do conclude in this or that *particular instance*, that so and so is wise and good, though we do not perceive its wisdom and goodness, but found our conviction solely on its being the divine will. But then, this is from our general *conviction* that God *is* wise and good; not from our attaching no meaning to the words wise and good, except the divine will. . . . And so it is in many other cases. You have read (suppose) several works of a certain author, and have found them all highly interesting and instructive. If, then, you hear of his bringing out a new work, you expect, before you have seen it, that it will be a valuable one. But this is not from your meaning by a "valuable work" nothing at all but that it comes from his pen.'[2]

'It is true . . . that we are *commanded* to do what is right, and forbidden to do what is wrong,' but 'it is not true that this is the only meaning of the words "right" and "wrong." And it is true that God *will* reward and punish', but not 'that a calculation of reward and punishment constitutes the whole notion of Duty.'[3] Paley, in short, identifies goodness with characteristics which are merely its invariable and necessary accompaniments.

The 'argument from trivialization' was a favourite one with Whately. It occurs in his *Lessons on Morals*, which appeared in 1855, four years before the edition of Paley; and may be found also in a letter written a year before that, in which he

[1] Whately's *Paley*, p. 24. [2] p. 25. [3] p. 27.

says, referring to those who say that 'right' means commanded by God:

'One might ask one of these moral teachers, "Do you think it right to obey the Divine will?" . . . do you think that God has a just claim on your obedience? For, if you do, then to say that it is "morally right" to obey Him, and yet that all our notions of morality are derived from our notions of His will, is just to say that what He has commanded is—what He has commanded!'[1]

Paley could, of course, have avoided any charge of circularity by simply denying that there is any meaning in such assertions as that God is just, or that it is right to obey Him. He could have said that since God has power to enforce His commands, the fact that it is merely tautological to call them just does not matter. And indeed he comes very close to saying precisely this when he tells us that the appeal to self-interest takes us to 'the bottom of the subject', and that 'no further question can reasonably be asked'. But he is not quite consistent about this; the position he takes up is, in fact, rather remarkable. Immediately after having defined 'right' as 'consistency with the will of God', he himself raises the question, 'But if the divine will determine the distinction of right and wrong, what else is it but an identical proposition to say of God, that He acts *right*? or how is it possible to conceive even that He should act wrong? Yet these assertions are intelligible and significant.' Archbishop Whately, or Professor Moore, could not have said more; how, then, does Paley escape?

'The case', he says, 'is thus: By virtue of the two principles, that God wills the happiness of his creatures, and that the will of God is the measure of right and wrong, we arrive at certain conclusions; which conclusions become rules; and we soon learn to pronounce actions right or wrong, according as they agree or disagree with our rules, without looking any further; and when the habit is once established of stopping at the rules, we can go

[1] *Life of Archbishop Whately*, vol. ii, p. 314.

back and compare with these rules the divine conduct itself: and yet it may be true (only not observed by us at the time) that the rules themselves are deduced from the divine will.'

'Arguing in a *circle*', Whately comments on this, 'is very common; with crafty sophists, from design, and with bad reasoners, from confusion of thought. But the former are very careful to conceal the fallacy; and the latter do not perceive it. It is very strange that Paley should perceive and acknowledge that he is involved in a circle, and should yet adhere to it.'[1] There does not seem to be anything that one could add to this or take away from it. And yet—*has* Paley argued in a circle here? Paley was, in his time, a Cambridge philosopher; and had he been a Cambridge philosopher in our time he might have answered his own question in some such way as this: 'We can intelligibly ask whether what God does and commands is right, and we can intelligibly ask whether what produces happiness is right. But this does not mean that in each case we are asking whether the subject possesses some possibly "non-natural" predicate distinct from both "conforming to God's will" and "productive of happiness". There is no such thing as *the* meaning of "right". The acts which we have learnt to describe so are in fact both done and commanded by God, and productive of happiness. And when we ask whether what God wills is right, we are asking whether all God's deeds and commands are like these ones in promoting happiness; while when we ask whether promoting happiness is good, we are asking whether all felicific actions are like these ones in being done or commanded by God.' And is this so very different from what Paley actually said?

At all events, what Paley said at this point was in part an unconscious prophecy. His *Moral and Political Philosophy* first appeared a few years after Bentham's *Principles of Morals and Legislation*; but Paley crystallized the theological

[1] p. 88.

Utilitarianism of the preceding period, while Bentham's secular Utilitarianism caught the ear of the age which followed it. And whereas the older school had defined virtue as obedience to the will of God, and made the promotion of happiness its 'subject-matter', in Bentham the latter became the definition.

> 'Of an action that is conformable to the principle of utility one may always say either that it is one that ought to be done, or at least that it is not one that ought not to be done. One may say also that it is right it should be done; at least that it is not wrong it should be done. . . . When thus interpreted, the words *ought*, and *right* and *wrong*, and others of that stamp, have a meaning; when otherwise they have none.'[1]

But he goes on immediately to treat the 'principle of utility'—that is, 'that principle which approves or disapproves of any action whatsoever, according to the tendency which it appears to have . . . to promote or oppose happiness'[2]—not as a mere definition, but as the self-evident premiss of all true and significant moral propositions. 'Is it susceptible of direct proof? it should seem not: for that which is used to prove everything else, cannot itself be proved. . . . To give such proof is as impossible as it is needless.'[3] And Bentham could not have given Paley's answer to the charge that he was guilty of the naturalistic fallacy at this point; for in him, although the promotion of happiness takes the place of obedience to God's will as the definition of right action, obedience to God's will does not take the place of promoting happiness as the 'subject-matter' of it, but simply disappears. Bentham has not, that is to say, a definition and a rule distinct from the definition which may, through a natural shift of usage, change places with it; in him the definition and the rule are one.

So far as I am aware, the first writer to charge Bentham, in effect, with committing the naturalistic fallacy was Sidgwick, though he found it hard to believe that Bentham's

[1] *Principles of Morals and Legislation*, I. x. [2] I. ii. [3] I. xi.

identification of his rule and his definition was seriously intended. His criticism occurs in a brief footnote,[1] where he argues that

> 'when Bentham explains . . . that his fundamental principle "states the greatest happiness of all those whose interest is in question as being the right and proper end of human action", we cannot understand him really to *mean* by the word "right" "conducive to the general happiness", though his language in other passages of the same chapter . . . would seem to imply this; for the proposition that it is conducive to general happiness to take general happiness as an end of action, though not exactly a tautology, can hardly serve as the fundamental principle of a moral system.' [It is not 'exactly' a tautology because to aim at some end is not necessarily the best way of actually realising it.]

This note is given simply as illustrating the point that the description of something as 'right' cannot always mean merely that it is the fittest means to some end, because we sometimes 'regard as "right" the adoption of certain ends—such as the common good of society, or the general happiness'. It is a note of considerable historical importance, as there is good reason to believe that it inspired Professor Moore's work on the 'naturalistic fallacy'. It is cited in a section in which Professor Moore begins by saying that, so far as he knows, 'there is only one ethical writer, Prof. Henry Sidgwick, who has clearly recognized and stated' the fact that 'good' is indefinable.[2]

Sidgwick would certainly have been the last to have claimed any originality for himself at this point. In his *History* the first point which he notices in his account of Price is 'his conception of "right" and "wrong" as "single ideas" incapable of definition or analysis'. (I suspect that 'single' here is a misprint for 'simple'.) Nor, I should say, did he imagine that he was original in his use of the 'argument from trivialization', apart from his being the first (if he was the

[1] To *The Methods of Ethics*, I. iii. 1. [2] *Principia Ethica*, p. 17.

first) to apply it to the particular case of Bentham. He was familiar with the work of Shaftesbury, Hutcheson, and Price, and would know the passage from Shaftesbury not only directly, but also as quoted against Paley in Dugald Stewart's *Philosophy of the Active and Moral Powers of Man*,[1] a work which he regarded as 'a lucid, comprehensive, and judicious attempt to put together the elements of truth in the work of preceding writers, including Shaftesbury and Adam Smith, into a harmonious and coherent statement of the results of impartial reflection on the moral consciousness'.[2] (Three writers of the period before Sidgwick's—Price, Stewart, and Whately—have now been mentioned as using the argument from trivialization against Paley; and it seems not unlikely that others did so too, particularly since Paley himself admitted that it might be raised.)

Sidgwick came still closer to the language of Professor Moore in a work published posthumously only a year before the first appearance of the *Principia Ethica*, on *The Ethics of Green, Spencer, and Martineau*. In the second lecture on Spencer, the latter's contention that '"pleasurable" and "painful"are the primary meanings of "good" and "bad"' is met with the observation that

'we must distinguish inquiry into the meaning of words from inquiry into ethical principles. I agree with Mr. Spencer in holding that "pleasure is the ultimate good", but not in the meaning which he gives to the word "good". Indeed, if "good" (substantive) means "pleasure", the proposition just stated would be a tautology, and a tautology cannot be an ethical principle.'[3]

There is, in fact, a far-reaching similarity in aim, or shall we say in provocation, between this work of Sidgwick's and the *Principia Ethica*, as both books attempt to show that the evolutionary ethical naturalism of Spencer and the 'metaphysical' ethics of T. H. Green suffer from a common error.

[1] II. v. 1. [2] *History of Ethics*, ch. iv.
[3] *Ethics of Green*, &c., p. 145.

Professor Moore identifies this common error with the 'naturalistic fallacy'; but although Sidgwick, as we have just seen, does mention this in connexion with Spencer, he treats it, as we have earlier suggested it ought to be treated, as an element in a larger error, namely, the denial of the autonomy of Ethics.

'Spencer and Green', he says in his opening paragraph, 'represent two lines of thought divergent from my own in opposite directions, but agreeing in that they do not treat Ethics as a subject that can stand alone. Spencer bases it on Science, Green on Metaphysics. In discussing Spencer', he goes on, 'we shall be dealing with an attempt to "establish Ethics on a scientific basis". Now this, I hold, cannot be done to the extent and in the manner in which Mr. Spencer tries to do it. "Science" relates to what is, has been and will be, Ethics to what ought to be; therefore the fundamental principles of the latter must be independent of the former, however important and even indispensable Science—especially Biology and Sociology—may be in the working out of the system of rules. And Science—in particular Psychology and Sociology—may trace the origin of moral sentiments and ideas, but it cannot itself supply a criterion of the validity of moral principles, or authority of moral sentiments.'

With this last point, as it has been developed both by Sidgwick and by other writers, we have already dealt quite fully. Nor need any more be said in order to establish the fact that Professor Moore's achievement has not been to work a revolution in Moral Philosophy, but simply to help keep alive, in our own age, the eighteenth-century tradition of sanity and logical rigour which Sidgwick (with Huxley the agnostic beside him and Whately the Archbishop behind him) kept alive in his.