

M I N D

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY



I.—G. E. MOORE'S LATEST PUBLISHED VIEWS ON ETHICS

BY C. D. BROAD

THE first six essays in the book *The Philosophy of G. E. Moore*, published in 1942 as Vol. IV in *The Library of Living Philosophers*, are devoted to Moore's ethical theories ; and Moore's comments upon them occupy the first ninety-three pages of his terminal essay. I suppose that this part of the terminal essay must contain Moore's latest published pronouncements on ethical problems. As such, it is of considerable interest and importance. Of the six ethical essays and Moore's comments on them I propose to select three for discussion here, *viz.* those of Frankena, Stevenson, and myself. Between them they cover the following four main topics, *viz.* (I) The distinction between 'natural' and 'non-natural' characteristics, (II) The 'autobiographical' analysis of moral indicatives, (III) The interconnections of value and obligation, and (IV) Ethical egoism and ethical neutralism. I propose to treat each of these topics in turn.

(I) *The distinction between 'natural' and 'non-natural' characteristics*

It is a well known doctrine of Moore's that the word 'good', in one important sense of it, stands for a characteristic of a peculiar kind which he terms 'non-natural'. In *Principia Ethica* he gave certain criteria for distinguishing 'natural' and 'non-natural' characteristics. The two marks of a *natural* characteristic were said to be (i) that it 'can exist in time all by itself', and (ii) that it is a 'part' of anything that it characterizes. I tried to show in my essay that these criteria are utterly unsatisfactory. Moore accepted that criticism ; and so we may henceforth regard that part of his doctrine as withdrawn.

© Thomas Nelson and Sons Ltd 1961

In my essay I suggested that Moore was almost certainly intending to deal with the same distinction (though he does not use the words 'natural' and 'non-natural') in the paper entitled "The Conception of Intrinsic Value" in his *Philosophical Studies* (1922). I understood his doctrine there to be as follows. (1) The characteristics of a thing may be divided into (a) those that *do*, and (b) those that do *not*, 'depend solely on its intrinsic nature'. (2) Those characteristics which do depend solely on the intrinsic nature of that which they characterize may be subdivided into (α) those which *are*, and (β) those which are *not* 'intrinsic'. (3) The *non-natural* characteristics of a thing are the members of the sub-class (a, β), *i.e.* those which *are* dependent solely on its intrinsic nature but are *not* intrinsic. The *natural* characteristics of a thing are the members of class (b) and the members of sub-class (a, α), *i.e.* they are those characteristics of it which *either* do not depend solely on its intrinsic nature *or* which depend solely on its intrinsic nature and are also intrinsic.

In his terminal essay Moore points out where I was right and where I was wrong in my interpretation of his doctrine in "The Conception of Intrinsic Value." I was right in thinking that he was concerned there with the distinction which he described in *Principia Ethica* by the words 'natural' and 'non-natural'. But I was wrong in thinking that he would admit there to be such a class of characteristics as (a, β), *i.e.*, ones which *do* depend solely on the intrinsic nature of that which they characterize and yet are *not* intrinsic. Moore says that he held that *all* characteristics which depend solely on the intrinsic nature of that which they characterize *are intrinsic*. And he held that goodness, in the fundamental sense in which he is here concerned with it, is intrinsic.

He thinks that my mistake may have arisen from the very unfortunate terminology which he used in "The Conception of Intrinsic Value." He admits that he there used the term 'intrinsic property' in such a way that there would be no inconsistency between the following three statements, (i) 'P is intrinsic', (ii) 'P is a property', and (iii) 'P is not an intrinsic property'. For, he says, his doctrine was that goodness (in the sense in question) *is* intrinsic and *is* a property and yet *is not* an intrinsic property of a good thing.

In view of this, I think that my misunderstanding was not only excusable but also fortunate, for it gave Moore an opportunity to remove what must have been a constant source of confusion even to wary readers. Henceforth, he says in the terminal essay, he will drop this terminology. In future, if I understand him aright,

he would call *all* those properties and *only* those properties of a thing, which depend solely on its intrinsic nature, 'intrinsic properties' of it. He would then sub-divide the intrinsic properties of a thing into 'natural' and 'non-natural'. And he would hold that goodness (in the sense in question) is a non-natural intrinsic property of a good thing. It will be noted that 'being an intrinsic property of a thing' is defined in terms of the notion of 'depending solely on the intrinsic nature of a thing'. The latter notion is elaborately expounded in "The Conception of Intrinsic Value". I did not criticize it in my essay, and Moore takes it for granted in his terminal essay; so I shall not discuss it here.

The verbal confusion is now removed, but we are left with the substantial question: What is Moore's criterion for distinguishing between those intrinsic properties of a thing which are *natural* and those which are *non-natural*? In "The Conception of Intrinsic Value" Moore gave two criteria, and the first of these may be subdivided into two complementary parts. In the amended terminology they may be stated as follows. (1.1) A complete enumeration of the *natural* intrinsic characteristics of a thing would be a *complete description* of that thing. (1.2) An enumeration which omitted any *natural* intrinsic characteristic of a thing would be an *incomplete description* of that thing. (2) The *natural* intrinsic characteristics of a thing seem to contribute towards describing its intrinsic nature in a way in which predicates of value do not.

In my essay I confined myself to (1.1), and said nothing about (1.2) or about (2). Moore admits in the terminal essay that (1.2) cannot be maintained as it stands. Suppose that P and π are two properties, *e.g.* being red and being coloured, such that anything that had P would, as a necessary consequence, have π . Then a description which included P would not be made incomplete merely by omitting π . And yet π might be a *natural* intrinsic characteristic. So (1.2) would have to be amended to run somewhat as follows: No description of a thing would be complete, if it omitted any *natural* intrinsic characteristic of it which is not conveyed by some one or some combination of its other natural characteristics. (I use the phrase 'P conveys Q' to mean the same as 'If anything had P, it would necessarily follow from that alone that it would have Q'.)

Moore admits in the terminal essay that he did not clearly distinguish criteria (1.1) and (1.2), on the one hand, from criterion (2), on the other. He says that he is now inclined to rely mainly on the following amended form of (2), *viz.* that, in one sense of

'describe', the mention of *any natural* characteristic of a thing contributes to some extent to describe that thing; whilst the mention of its *non-natural* intrinsic characteristics does not, in that sense, describe it at all. He admits that this is extremely vague, unless we can give some more definite information as to the particular sense of 'describe' which is here relevant.

I think it is fair to conclude that Moore, at the time when he wrote this terminal essay, was unable to give any satisfactory definition of, or criterion for, a 'non-natural characteristic'. But I think that we can go further. His suggested criterion, with its admitted vagueness, due to the uncertainty of the relevant sense of 'describe', is surely grist to the mill of supporters of what I will call 'non-predicative interpretations of moral sentences in the indicative'. If, as that theory holds, the word 'good' is not the name of a characteristic at all, but its use is, *e.g.* to express or to evoke certain emotions, then to call a thing 'good' would *not* contribute in any way to the description of it. And yet, owing to the likeness of grammatical form between such sentences as 'That is a pleasant emotion' and 'That is a morally good emotion', *e.g.* there might well seem to be something paradoxical in saying that the former did, and the latter did not, contribute towards describing the emotion. So one could understand why those who never questioned that moral sentences in the indicative assign a predicate to a subject, should sum up the situation by alleging that the word 'good' stands for a property of a peculiar kind, which does not contribute to describe its subject in the familiar way in which *e.g.* the property denoted by 'pleasant' does.

There remain two small points which are worth mentioning before leaving this part of the subject. (1) Moore says that, in his opinion, there are at least two kinds of intrinsic value, *viz.* goodness (in the sense in question) and beauty. But he does not hold, and never has held, that goodness, in that sense, is a determinable in W. E. Johnson's usage of that word. I must say, for my own part, that I should need a great deal of persuasion before I would admit that there is even a *prima facie* case for regarding beauty, in any sense of that word, as a form of *intrinsic* value.

(2) The other point is this. In the course of my essay I used an argument which presupposes that the pleasantness of a pleasant experience is dependent solely on its intrinsic nature. I assumed, *e.g.* that, in the case of a pleasant sensation, its pleasantness is always conveyed by some intrinsic pleasant-making sensible quality of it, such as its sweetness. Now Moore points out that the relation between the pleasant-making characteristics of an

experience and its pleasantness is almost certainly *not* that of conveyance, but is that of *causal determination*. I fully agree with that contention, and I will proceed to develop it in my own way, in which Moore might not have been willing to follow.

The essential point is that it is perfectly conceivable that two persons, or the same person on different occasions, should have sensations which were exactly alike in all their sensible qualities, and yet that one of them should be a pleasant experience and the other an unpleasant one. It is a very well founded empirical generalization, *e.g.* that the vast majority of human beings, whenever they have a sensation of the 'toothachy' kind, dislike that sensation for its characteristic sensible qualities. That is why we call toothachy sensations 'unpleasant'. But there is no kind of necessity about that generalization. It is perfectly conceivable that there might be persons who, when they had a sensation of precisely the same kind, always, or on certain special occasions, *liked* that sensation for those very same sensible qualities for which most persons at most times *dislike* such sensations. For any such person, on any such occasion, a toothache would be a *pleasant* experience. I would suggest, then, that the words 'pleasant' and 'unpleasant', as applied to experiences, often imply a well-founded empirical generalization, to the effect that the vast majority of people, on the vast majority of occasions when they have an experience of a certain kind, would like it (or dislike it, as the case may be) for its characteristic experiential qualities. But there is also, plainly, a non-statistical sense of the words 'pleasant' and 'unpleasant'. To call an experience 'pleasant' (or to call it 'unpleasant'), in this latter sense, means that the particular person, who has it on a particular occasion, then and there *likes* it (or *dislikes* it, as the case may be) for its characteristic experiential qualities. There is no kind of contradiction in saying that a particular experience, which would correctly be called 'pleasant' (or be called 'unpleasant') in the *statistical* sense, occurring on a particular occasion in a particular person, might be correctly called 'unpleasant' (or 'pleasant', as the case may be) in the *non-statistical* sense.

(II) *The 'autobiographical' Analysis of Moral Indicatives*

Consider the sentence: 'It was right for Brutus to stab Caesar', uttered at a certain moment by a person who is really considering what he is saying and is not merely talking like a parrot or giving an example in an essay. What I call the 'autobiographical' analysis of this sentence is, on its positive side,

that the speaker is intending to state, beside the historical proposition that Brutus stabbed Caesar, the autobiographical proposition that he himself is feeling a certain kind of emotion (*viz.* one of moral approval) in contemplating that historical proposition. On the negative side the theory is that the speaker is not intending to state anything else beside that historical and that autobiographical proposition and anything that may be logically entailed by them.

This must be carefully distinguished from what I have called 'the non-predicative theory of moral indicatives'. That holds that the speaker is stating nothing but the historical proposition; but that he would not have used the moral-indicative form of expression unless he were feeling moral approval towards it himself or had wanted to induce that emotion in his hearers. (The theory can, of course, take other forms, with something else substituted for 'moral approval'.) None of the essayists explicitly defends the non-predicative theory. But Professor Stevenson defends the autobiographical analysis against certain arguments which Moore had used in his paper "The Nature of Moral Philosophy" in *Philosophical Studies*. Moore, in his reply, says that he would be more inclined to accept the non-predicative theory than the autobiographical analysis, if he were to accept either.

Before going further it is worth while to note that the autobiographical analysis might take two different forms, which I will call 'occurrent' and 'dispositional'. On the occurrent form of it, a person who says at a certain moment that X is right is saying that he is at that moment feeling moral approval for X. On the dispositional form of the theory, he is saying that he is generally disposed to feel moral approval when he contemplates actions like X. Moore distinguished those two forms of the theory in his paper "The Nature of Moral Philosophy". But Stevenson considered only the occurrent form, and therefore Moore also confines himself to that in his reply. This seems to me unfortunate, because the dispositional form is much more plausible than the occurrent form.

There is a matter, which seems to me quite simple, about which both Stevenson and Moore make terribly heavy weather. The essential point at issue can be put as follows. Suppose that the occurrent form of the autobiographical analysis were correct. Then A's utterance at *t* of a token of the type-sentence 'X is right' would be equivalent in meaning to his uttering a token of the type-sentence 'I am now feeling moral approval of X'. Similar remarks apply, *mutatis mutandis*, to 'X is wrong'. Now

that makes the predicates 'right' and 'wrong' to be doubly relational, for it makes them involve a relation to a speaker and to a time. It follows that 'right' would have a systematically different meaning on every different occasion on which it is predicated, even by the same person, beside having a systematically different meaning corresponding to each different person who predicates it on any occasion. Now the word 'right' seems *prima facie* not to answer to those conditions. It seems to be used as if it could stand for precisely the same characteristic when predicated by different persons or on different occasions by the same person. Moore's arguments against the occurrent autobiographical analysis in "The Nature of Moral Philosophy" are simply various ways of trying to exhibit strikingly certain aspects of this *prima facie* conflict between the common usage of the word 'right' and the usage which would seem to be required if the occurrent autobiographical analysis were correct.

One of Moore's arguments was concerned with the possible alteration in a person's emotional attitude towards the same action, if he should contemplate it on successive occasions. Stevenson's criticism of this argument brings out an important point about the use of tenses in such sentences as 'X is right', 'X was right', and 'X will be right'. The point may be put as follows. Suppose that A says at *t* 'I now approve of X, but I formerly contemplated it with disapproval'. Obviously his statement may be true. Now Moore has argued that, if the occurrent autobiographical analysis be correct, A's statement would be equivalent to 'X is now right, but was formerly wrong'. And he had pointed out that it is nonsensical to say, of *one and the same* action, that it was right at one time and became wrong later.

Now Stevenson quite justifiably challenges Moore's right to assert that the theory entails the equivalence mentioned above. Stevenson insists that the correct interpretation of the theory is as follows. If a person says 'X is right', he means that he is *now* feeling approval towards X, which is *now* being performed. If he says 'X was right', he means that he is *now* feeling approval towards X, which *has been* performed. The tense in the moral indicative refers only to the date of the action which is said to be right or to be wrong; and the principal tense in the autobiographical equivalent of that indicative is *always the present*.

If we accept this contention of Stevenson's, what really does follow from the autobiographical analysis, together with the fact of the change in A's attitude, is this. A can now correctly and truly say 'X was right'; and he could, at some former time, have said with equal correctness and equal truth 'X is (or was)

wrong'. But at no time could he correctly and truly say 'X was right at one time and is now wrong'. For that would be equivalent to uttering the sentence 'I now approve of X, which happened in the past, and disapprove of X, which is happening now'. This is doubly nonsensical, since it asserts that the speaker had, at the same time, incompatible emotional attitudes towards one and the same particular action, and it implies that one and the same action was done at two different times.

Another argument in Moore's paper on "The Nature of Moral Philosophy" may be put as follows. Suppose that A and B contemplate the same act X at the same time *t*. A may say 'I approve of X', and B may say 'I disapprove of X', and both may be telling the truth. Now, if the analysis under discussion be correct, A's statement is equivalent to his saying, 'X is right', and B's statement is equivalent to his saying 'X is wrong'. Now the two latter statements conflict logically, whilst the two former are logically compatible. Therefore they cannot be equivalent each to each.

The true account of this situation is admirably brought out by Moore in his terminal essay. It is this. If the analysis under discussion be admissible, A can correctly and truly say 'X is right', and B can at the same time correctly and truly say 'X is wrong'. But no-one at any time can correctly and truly say 'X is both right and wrong'. For anyone who did so would, according to the proposed analysis, be saying 'I now approve and disapprove of X'. Now that could not be truly said by A, who approves and does not disapprove of X; nor by B, who disapproves and does not approve of X; nor by any third person, since no-one can entertain simultaneously incompatible emotional attitudes towards the same object.

This amendment, however, does nothing to diminish the force of Moore's original argument against the occurrent autobiographical analysis, *viz.*, that, according to it, A and B do not differ in opinion when one of them pronounces an action to be right and the other pronounces the very same action to be wrong. This is recognized by Stevenson, who proceeds to meet it by making two additions to the proposed analysis.

The first is to point out that, although A and B would not differ in opinion, in the sense of holding incompatible beliefs, they would do so in the wider sense of having opposed emotional attitudes toward the same object. The second is to remind us that in such situations each person would generally seek to alter the emotional attitude of the other and make it resemble his own.

Stevenson admits that, even when due weight has been given to these two considerations, the occurrent autobiographical analysis is not wholly satisfactory. Suppose that A asks B 'Is X right?' A is not as a rule wanting to find out whether *he himself* now approves of X, but whether B or most other people would do so. Or, again, A may disapprove of X and may know that B approves of it, and the motive of his question may be to induce B to change his attitude. Lastly, if A asks *himself* 'Is X right?', he is certainly not trying to find out whether he now approves of X. The situation probably is that he has conflicting attitudes towards X, in respect of various aspects of it, and that he is seeking to straighten them out.

Moore does not seriously dispute anything that Stevenson here says. He tells us that he has always recognized that difference of 'opinion' covers opposition of emotional attitude, but that he used not to think it possible that moral conflicts could be merely of that kind. He is now inclined to think that moral disagreement *may* be nothing but opposition of emotional attitude; but he is also inclined about equally strongly to think that it involves a logical conflict between incompatible beliefs. Stevenson, he says, has given no reasons for his own alternative; he has merely shown that certain arguments against it are inconclusive. If Moore felt obliged to abandon his own theory, he would not be inclined to stop at the stage of the occurrent autobiographical analysis, but would prefer to accept some form of *non-predicative* theory. Moore says that he is, in fact, now quite strongly disposed to think that, when a person utters the sentence 'X is right', he is not asserting *anything* that could be true or false, not even the autobiographical proposition that he now approves of X. But Moore says that he also continues to have some inclination to hold his old view. And he cannot say which of these inclinations is the stronger.

(III) *The interconnections of Value and Obligation*

The longest and most complex essay in the ethical part of the book is that of Professor Frankena, and the part of Moore's reply which deals with it is also highly involved. The question at issue is the connection between the fact that a state of affairs would be *intrinsically good* and a person's being under an *obligation* to seek to *bring it into existence*. Moore had made certain statements on this topic in his various ethical writings, and Frankena discusses their truth and their compatibility with

Moore's characteristic doctrines that good is a simple, indefinable, intrinsic, and non-natural characteristic.

The best way to convey an idea of the discussion is to take in turn the points which Frankena enumerates in the summary at the end of his essay, and to consider, in each case, Moore's treatment of them.

Point 1. This divides into two propositions, which I will call (1, *a*) and (1, *b*). The former is the contention that, if good (in Moore's sense) be *simple*, then the statement 'I am morally bound to do Y' cannot *mean* the same as the statement 'Y will produce more good or less evil than any other act open to me'. The latter is the contention that the same negative consequence follows from the supposition that good (in Moore's sense) is *intrinsic*, in the sense explained by him.

(1, *a*) After a good deal of discussion on alleged obscurities and ambiguities in Frankena's reasoning, Moore proceeds to state formally what he takes to be Frankena's argument on this point. I have very little doubt that this is a correct account of what was present, in a less precise form, in Frankena's mind, and so I shall adopt it. The argument may be stated as follows. The proposition that good is *simple* entails that statements of the form 'X is good' neither include nor are identical with statements of obligation. That entails that statements of the form 'X is good' are *not normative*. That in turn entails that statements of the form 'Y will produce the most good or the least evil of all the acts open to me' are *not normative*. And that entails that statements of the latter form are not identical in meaning with statements of the form 'I ought to do Y'.

Now Moore holds that the fundamental step in this argument is the second, *viz.* that if statements of the form 'X is good' neither include nor are identical in meaning with statements of obligation it follows that such statements are not normative. The validity or invalidity of this step depends on what Frankena means by 'normative', and that (Moore alleges) is not made perfectly clear in his essay. But, setting aside minor verbal inconsistencies, it seems fairly plain that what he intends is the following. S is a normative statement about an action, if and only if it follows from the very nature of that statement that that action ought to be done. If we accept this account of 'normative', we see that the transition in step 2 depends on the tacit assumption that nothing can follow from the very nature of a statement except what is identical with or is a part of what is meant by the latter. There is in fact no doubt that Frankena does assume this premiss, for elsewhere in his essay he makes it

quite explicit that he thinks that the two propositions 'Q follows from the very nature of P' and 'Q is *synthetically, though necessarily*, connected with P' are mutually exclusive. Now Moore rejects this premiss, and therefore sees no reason to accept step 2 of Frankena's argument.

As this is an important point, I will state all that is to be found in Moore's terminal essay on this topic. In the first place, he gives an example taken from Professor Langford's essay in the same volume. He says that, in his opinion, it *does* follow, from the very nature of the statement 'This is a cube.', that this has twelve edges; whilst the latter is *not* identical with nor a part of the meaning of the former. Secondly, in another part of his essay, Moore makes the following general assertions. He says that he uses the phrase 'Q follows from P' to mean that the conjunction 'P & not-Q' is *self-contradictory*. But he holds that such a conjunction may be self-contradictory *without* 'Q follows from P' being *analytic*.

If we put all this together, we see that what Moore is maintaining is the following. Even though good be *simple*, the conjunction 'I ought to do Y, and Y will *not* produce as good consequences as some other action open to me' may be *self-contradictory*, in that sense (whatever it may be) in which the conjunction 'This is a cube, and this has *not* twelve edges' is self-contradictory. I should agree that this is quite possible, provided that *ought* itself is not simple, but contains *good* in its analysis. But, if *good* and *ought* were both simple, I cannot for the life of me see how the conjunction in question could be *self-contradictory*, in any generally accepted sense of that phrase. It might, however, be *self-evidently impossible*, without being self-contradictory in the formal sense, if we admit the possibility of necessary connections and disconnections which are synthetic, but obvious on inspection. I should add, perhaps, that I am extremely doubtful whether the conjunction 'This is a cube, and this has not twelve edges' *is* self-contradictory. I should suspect that what is so is the conjunction of this conjunction with certain of the axioms of three-dimensional Euclidean geometry. If so, it is not very helpful as an analogy to the ethical propositions under consideration.

Moore remarks that Frankena might reply to his criticisms on step 2 of the argument by saying that he uses the word 'include' in such a way that 'Q is included in P' covers *inter alia* 'Q follows necessarily but synthetically from P'. But that would not help Frankena's argument, since it would save step 2 only at the expense of step 1. For, if 'include' be used in this

extended sense, there is no reason why the simplicity of *good* should prevent statements of the form 'X is good' from 'including' statements of obligation.

(1, b) This is the contention that, if good be *intrinsic* in Moore's sense of that word, then the statement 'I am morally bound to do Y' cannot mean what is meant by 'Y will produce more good or less evil than any other act open to me'.

Moore says that the argument is precisely the same as that in (1, a), with 'intrinsic' substituted for 'simple'. It therefore suffers from the same defect, *viz.* that the second step is unjustified, for the reasons given above. But it suffers from a further defect. For the first step, which was quite legitimate in (1, a), ceases to be so when 'intrinsic' is substituted for 'simple'. From the hypothesis that good is *intrinsic*, in Moore's sense, it would *not* follow that statements of the form 'X is good' neither include nor are identical with statements of obligation.

In order to discuss this, we must remember what Moore does and what he does not mean by calling a characteristic 'intrinsic'. To say that P is an intrinsic characteristic of X means that the possession of P by X depends solely on X's intrinsic nature. Now, in the first place, it does *not* follow from this definition that every intrinsic characteristic of X must be a *pure quality*. No doubt, if goodness were a pure quality, whether intrinsic or not, it would follow at once that 'X is good' could not be identical with or include any statement of obligation. For the latter would involve *relations* to an actual or possible agent. But Moore has distinguished between the 'external' and the 'internal' relational properties of a thing; and, whilst no *external* relational property of a thing could be intrinsic, there is nothing to prevent its *internal* relational properties from being so.

We may put the matter as follows. We must distinguish between what we might call 'categorical' and 'conditional' relational properties, though Moore does not use those terms. It would be a *categorical* relational property of a certain bit of arsenic to be poisoning Mr. Jones at a certain moment. That property would be external and non-intrinsic; for that bit of arsenic would not be having it unless Mr. Jones had existed and had swallowed it. It is a *conditional* relational property of any bit of arsenic to be poisonous, *i.e.* to be such that it *would* poison a man, *if* there were one and *if* he were to swallow it. This property, though relational, may be internal and intrinsic; for a bit of arsenic would have it even though there had never been any men or though no man had ever swallowed it. Similarly, if goodness be an *intrinsic* property of X, the statement 'X is good'

cannot include or be identical in meaning with any such *categorical* statement as 'Z ought to desire X' or 'Z ought to try to produce X'. But there is nothing to prevent its including or being identical with some *conditional* proposition of the form 'If there were a person who fulfilled such and such conditions, he would be under an obligation to desire X or to try to produce X'. For X could have that property, even if there had never been any persons, or if no person had ever fulfilled the required conditions.

Whilst I admit the validity and the importance of the distinction which Moore draws here, I do think that it is rather misleading to say of even a *conditional* relational property that it 'depends solely on the intrinsic nature of its possessor'. Surely there is an important sense in which the poisonousness of arsenic depends just as much on the intrinsic nature of a *living organism* as on the intrinsic nature of arsenic. In the same sense and to the same degree the property of being such that, if there were a person and he were to fulfil certain conditions, he would be under an obligation to try to produce X, depends just as much on the intrinsic nature of *moral persons* as on that of X. No doubt arsenic would have been poisonous, even if there never had been and never will be any living organisms; but at least we can say that the very notion of poisonousness involves the notion of organisms and vital processes, and that no amount of reflection on arsenic in isolation could have supplied the latter notions.

Point 2. Frankena's second point really divides into seven interconnected propositions. It may be stated as follows. If value be either (a) *simple*, or (b) *intrinsic*, then it cannot be either (α) *normative*, or (β) *non-natural*, or (γ) *definable in terms of obligation*. And, that being so, (c) there is no reason to think that it is *incapable of being defined in non-ethical terms*.

It is evident that we thus have six hypothetical propositions, which arise by combining in turn each of the two antecedents (a) and (b) with each of the three consequents (α), (β), and (γ). In addition to these six hypotheticals there is the seventh proposition (c), which Frankena states in the form 'In that case there is no reason to regard value as being incapable of definition in non-ethical terms'. We may label the six hypotheticals as (2a, α), (2a, β), (2a, γ), and (2b, α), (2b, β), (2b, γ). The seventh proposition may be labelled (2c).

Moore claims to have dealt with (2a, α), (2a, γ), (2b, α), and (2b, γ) in his discussion of (1, a) and (1, b). He has admitted (2a, γ), i.e. that, if good be simple, it cannot be defined in terms of obligation, since it would be indefinable. He has rejected (2a, α),

(2b, α), and (2b, γ). It remains, therefore to deal with (2a, β), (2b, β), and (2c). That we will now proceed to do.

(2a, β). This is the proposition that, if good be *simple*, it cannot be *non-natural*. The essence of Frankena's argument is as follows. If good were simple, it would not be normative. If it were not normative, there would be no reason to think it non-natural. Therefore, if it were simple, it would be non-natural.

Now the first step has already been discussed and rejected. And, even if both it and the second step were accepted, the correct conclusion would be only that, if good were simple, there would be *no reason to think* that it is non-natural. There would be no justification for the stronger conclusion that it *would not be* non-natural.

(2b, β). This is the proposition that, if good be *intrinsic*, it cannot be *non-natural*. The argument is the same as before, with 'intrinsic' substituted throughout for 'simple'. The first step of this argument has already been discussed and rejected. And the argument has the same defect as (2a, β), *viz.* that of drawing a stronger conclusion than would be justified by its premisses, even if these were acceptable.

Before passing to (2c), it will be worth while to consider for a moment the second premiss, which is common to both the above arguments of Frankena's. This is the proposition that, unless good were *normative*, there would be no reason to think it *non-natural*. The essence of Frankena's contention on this topic is as follows. In his opinion, the main point of the doctrine that intrinsic value is non-natural is that it cannot be reduced to purely psychological, sociological, biological, or metaphysical terms. Now it seems to him that the only feature in moral judgments which can plausibly be held not to be so reducible is their ostensibly *normative* character, *i.e.* 'the fact that they seem to be saying of some agent that he *ought* to do something'. He concludes that, unless intrinsic value 'in itself possesses a normative character or obligatoriness', there is no reason to think that it is not essentially reducible to the terms enumerated above.

(2c). This proposition, which comes immediately after the six hypotheticals which we have now discussed, is stated in the very obscure sentence: 'In that case there is no reason to think that good is not definable in non-ethical terms.' We naturally ask: 'In *what* case?' In the context Frankena might mean *either* that if his six hypotheticals were true there would be no reason to think that good is not definable in non-ethical terms, *or* that if their three consequents were true there would be no reason to think this. Moore does not consider the first of these alternatives, but confines his attention to the second. This is the

proposition that, if good be neither normative nor non-natural nor definable in terms of obligation, then there is no reason to think it is not definable in non-ethical terms.

The phrase 'not definable in non-ethical terms' needs a certain amount of unpacking. It will be best to start from the beginning. Good might be either (1) indefinable, or (2) definable. If indefinable, it might be either (1.1) identical with some admittedly non-ethical simple notion, *e.g.* pleasant, or (1.2) not identical with any admittedly non-ethical notion. If definable, it might be either (2.1) definable in wholly non-ethical terms, or (2.2) definable only in terms which are wholly or partly ethical. We could lump together the two alternatives (1.1) and (2.1) under the heading 'wholly *expressible* in non-ethical terms'; and the two alternatives (1.2) and (2.2) under the heading 'not wholly *expressible* in non-ethical terms'.

Now Frankena has argued that, if good were *simple*, it would be neither normative nor non-natural nor definable in terms of obligation. And we have interpreted (2c) to mean that, if good were neither normative nor non-natural nor definable in terms of obligation, there would be no reason to think that it is not definable in non-ethical terms. Putting the two together, we see that Frankena is committed to the proposition that, if good be *simple*, there is no reason to think that it is *not definable in non-ethical terms*. But, obviously, if it be simple, it cannot be *definable* in any terms whatever. So, in order to make sense of the above proposition, we must assume that Frankena is using the phrase 'not definable in non-ethical terms' in a loose sense which is equivalent to my phrase 'not wholly expressible in non-ethical terms'. What he is asserting is, in fact, the following proposition. If good be neither normative nor non-natural nor definable in terms of obligation, there is no reason to think that it is not *either* (i) definable in wholly non-ethical terms, *or* (ii) identical with some simple admittedly non-ethical notion. If that is what Frankena means, it may be doubted whether (2c) is more than a tautology; for the only attempt which he makes to define 'non-natural' seems to identify it with 'not wholly expressible in certain enumerated non-ethical terms'.

Point 3. Frankena's third point may be put as follows. If good were either (a) *normative* or (b) *non-natural* or (c) *not wholly expressible in ethical terms*, then (α) it would be *definable in terms of obligation*, (β) it would *not be simple*, and (γ) it would *not be intrinsic*. The third point is therefore the conjunction of the nine hypothetical propositions which arise by uniting (a), (b), and (c) in turn as antecedents with (α), (β), and (γ) as consequents,

Now of these nine hypotheticals the following have already been dealt with. $(3a, \alpha)$ is the contrapositive of a step in the argument for point (1, a), which Moore has discussed and dismissed. $(3a, \beta)$, $(3a, \gamma)$, $(3b, \beta)$, and $(3b, \gamma)$ are the contrapositives of $(2a, \alpha)$, $(2b, \alpha)$, $(2a, \beta)$, and $(2b, \beta)$ respectively. And these have been discussed and rejected by Moore. Again, if $(3b, \alpha)$ be granted, then $(3c, \alpha)$ becomes superfluous. For it is admitted that, if good be not wholly expressible in non-ethical terms, it is non-natural. And, if this be combined with $(3b, \alpha)$, we can infer $(3c, \alpha)$. $(3b, \alpha)$ embodies Frankena's conviction, already discussed, that the only fundamentally ethical notion is that of obligation. We are thus left with only $(3c, \beta)$ and $(3c, \gamma)$. These are more simply expressed in the equivalent form of their contrapositives. If we do this, and combine them, they amount to the proposition that, if good were either *simple* or *intrinsic*, it would be *wholly expressible in non-ethical terms*, i.e. it would be *natural*. This will best be treated incidentally in connection with the remaining points in Frankena's summary.

Point 4. Frankena's fourth point is that Moore has given no adequate reason for rejecting the view that 'good' is *definable in terms of obligation*.

Moore begins by admitting that he has given no *conclusive* reason. But he thinks that he can give *good* reasons. The gist of his argument is as follows. He considers three alleged definitions of 'good' in terms of obligation, which Frankena proposes. He rejects one of them on the ground that the two propositions suggested as definiens and definiendum do not even mutually entail each other. As regards the other two, he admits that there is mutual entailment between the definiendum and the suggested definiens. But he holds that that kind of logical relation can hold between two propositions without it being a case of two sentences with one and the same meaning. The test for the latter is to ask oneself the question 'Can I think of the one without *ipso facto* thinking of the other?' In each of these two cases he holds that that is possible, and therefore that there is not identity of meaning. Now, Moore says, he cannot think of any *other* plausible instances of mutual entailment between a value-proposition and an obligation-proposition. Therefore he holds that he has given sound, though not conclusive, reasons for thinking that goodness cannot be defined in terms of obligation.

I will now say something about the three proposed definitions. (1) 'X is intrinsically good' means what is meant by 'If one is capable of producing X, one has a *prima facie* duty (in Ross's sense) to do so'. This is rejected by Moore (rightly, I think), on

the ground that the former proposition might easily be true when the latter was false.

(2) I am going to formulate the second in a slightly modified form of Moore's interpretation of Frankena's rather vague statement. It will run as follows. 'X is intrinsically good' means what is meant by 'The mere fact (if it were a fact) that A could do Y and that Y would produce X would suffice to supply *some* reason for thinking that A *ought* to do Y'. Moore holds that these two propositions do entail each other. But he considers it obvious that a person could think of X as being intrinsically good, without *ipso facto* thinking of it as having this other complicated property, which is conveyed by and conveys its intrinsic goodness.

(3) Frankena quotes a certain alleged mutual entailment, which Moore gave in his *Ethics*, and asks why this should not be regarded as a definition. We need not trouble ourselves here about *this* particular alleged mutual entailment, because Moore says that he does not now think that it holds, or that the sentence quoted correctly expressed what he had in mind when he wrote his *Ethics*. Instead, we may confine our attention to the amended formula which he now proposes in its place. It runs as follows: 'X is intrinsically good' entails and is entailed by 'If an agent were a Creator, before the existence of any world; and if the only two alternatives open to him were (i) to create a world which consisted only of X, or (ii) to bring it about that there should never be a world at all; then it would be his duty to choose alternative (i), provided (a) that he knew for certain that these were the only two alternatives open to him, and (b) that he did not think it wrong to choose alternative (i)'. Moore says that it seems obvious to him (and who shall deny it?) that a person could think of the former proposition without *ipso facto* thinking of the latter.

Point 5. The main assertion in Frankena's fifth point, and the only one which Moore discusses, is the following. Frankena alleges that, even though it be *intrinsic value* which makes a thing such that it *ought* to be pursued or brought into being by a competent agent, still Moore has given no good reason why 'intrinsic value' might not be *definable in wholly non-ethical terms*.

What Frankena has in mind is no doubt this. He is alleging that there is no obvious reason why a *purely natural* characteristic, e.g. pleasantness, should not be such that the mere fact that a thing would have it would provide some ground for thinking that any agent, who could produce that thing, *ought* to do so.

In order to discuss this, let us begin by defining what Moore calls an 'ought-*implying* property', and what I prefer to call an 'ought-*inclining*' property. The sentence 'P is an ought-inclining property' is to mean what is meant by 'The mere fact that a thing would have P would suffice to provide *some* ground for thinking that any agent, who could bring such a thing into being, *ought* to do so'. Now Moore admits that *intrinsic goodness* is an ought-inclining property. He holds, moreover, that the intrinsic goodness of a thing always depends on the presence in it of some *natural* characteristic or other which is what I will call 'good-making'. Let Q be any *good-making* natural characteristic. Then anything that had Q would, of necessity, have intrinsic goodness. And the mere fact that anything had intrinsic goodness would suffice to provide some ground for thinking that any agent, who could bring such a thing into being, ought to do so. It follows at once that Q, though a *natural* characteristic, will also answer to the definition of an 'ought-inclining property'. And this can be generalized at once for every natural characteristic which is good-making.

There is no doubt, then, that Frankena is right in holding that there can be, and in fact are, *natural* characteristics which are ought-inclining. It is plain that he thinks that this fact entails that intrinsic goodness *either* (i) is *definable in terms of ought*, or (ii) is a *natural* characteristic. He thinks that, if the former alternative were fulfilled, there would be some reason to think that intrinsic goodness is *non-natural*. For, as we have seen, he regards 'ought' as *the* ethical notion *par excellence*, and as such the most plausible instance of a non-natural notion. On the other hand, he thinks that, if intrinsic goodness be *not* definable in terms of ought, then (in view of the fact that an ought-inclining property *can be natural*) there will be no valid reason for thinking that intrinsic goodness is non-natural.

Now Moore gives an argument which, he thinks, tends to show that intrinsic goodness cannot be identical with any natural property, even if it be not definable in terms of ought. The argument runs as follows.

Admittedly some *natural* characteristics are ought-inclining. But only *intrinsic* natural characteristics can be such. For a natural characteristic is ought-inclining only through being good-making. And only intrinsic natural characteristics convey intrinsic goodness. So the question reduces to whether intrinsic goodness could be identified with any *intrinsic* natural characteristic. After these preliminaries the argument continues as follows.

The number of ought-inclining intrinsic natural characteristics is, Moore asserts, certainly very great and possibly infinite. Plainly, we cannot identify intrinsic goodness with any particular one of them or with the aggregate of all of them. Moore thinks it obvious, moreover, that intrinsic goodness could not be identified with the *disjunction* of all these natural characteristics. Suppose that there were *one single* non-disjunctive intrinsic natural characteristic, which was (a) ought-inclining, and (b) was conveyed by each of the other ought-inclining natural characteristics. Then it might be plausible to identify intrinsic goodness with *it*. But there seems to be no one natural characteristic answering to these conditions. Therefore there does not appear to be any ought-inclining natural characteristic with which intrinsic goodness can be identified. And it certainly cannot be identified with any intrinsic natural characteristic which is *not* ought-inclining. Therefore it cannot be identified with any natural characteristic whatever.

I think that this argument is valid, so far as it goes. But it would not satisfy a person who might suggest that 'X has intrinsic goodness' means what is meant by 'X has *some* intrinsic natural characteristic or other which is ought-inclining'. I do not know whether Moore would count this as identifying intrinsic goodness with the *disjunction* of all ought-inclining intrinsic natural characteristics. I do not think that he would. But, if he did, I should be inclined to ask: What precisely is the objection to such an 'identification'? The advantages of the suggestion are that it avoids postulating *two* indefinable non-natural characteristics, and defines the less specifically ethical one ('intrinsically good') in terms of the more specifically ethical one ('ought'). The final objection would have to be that one can think of intrinsic goodness without *ipso facto* thinking of even so indeterminate a notion as that expressed by the phrase 'some intrinsic natural ought-inclining characteristic or other'. But is that really at all certain?

Moore's latest published Views on the Connection of Good, Better, and Ought. On pp. 606 to 611 of his terminal essay Moore formulates four very complicated pairs of mutually entailing propositions, which express the views, which he held at the date of writing, about the interconnections of 'good' and 'ought' and of 'better' and 'ought'. I am going to state them in my own way; but what I shall say is, I think, equivalent to what Moore had in mind and is perhaps somewhat easier to grasp.

I shall begin by defining certain statements. (1.1) 'P is a *good-making* characteristic' means what is meant by 'If X did

have or now has or will have P, it follows that X then was or now is or will then be intrinsically good ; and if X should have had or should now have or should be going to have P, it would follow that X would have been or would now be or would be going to be intrinsically good '. (1.2) 'P is a *bad-making* characteristic' is defined in a precisely similar way, with '*intrinsically bad*' substituted throughout for '*intrinsically good*'. (1.3) 'P is a *valifying* characteristic' means what is meant by 'P is either a good-making or a bad-making characteristic'. (2) 'P is *more strongly good-making or less strongly bad-making* than Q' means what is meant by 'If X did have or now has or will have P (and *no other* valifying characteristic), and if Y did have or now has or will have Q (and *no other* valifying characteristic), it follows that X then was or now is or will then be better than Y ; and if X should have had or should now have or should be going to have P (and *no other* valifying characteristic), and if Y should have had or should now have or should be going to have Q (and *no other* valifying characteristic), it would follow that X would have been or would now be or would be going to be intrinsically better than Y '.

We can now formulate the four pairs of mutually entailing propositions.

First Pair. (i) P is a *good-making* characteristic. (ii) If there had been, or in fact was, an agent who, before any world existed, (a) *knew* (α) that if he chose he could create a world characterized by P, (β) that he could so choose, and (γ) that if he did not so choose no world at all would ever exist ; and who (b) *did not believe* that this choice would be wrong ; then it would have been, or in fact was, the *duty* of that agent to make that choice.

Second Pair. (i) P is a *good-making* characteristic. (ii) P is an *ought-inclining* natural characteristic.

Third Pair. (i) P is a *more strongly good-making or a less strongly bad-making* characteristic than Q. (ii) If there had been or in fact was, an agent who, before any world existed, (a) *knew* (α) that if he chose he could create a world characterized by P, (β) that he could so choose, and (γ) that, unless he were so to choose, a world characterized by Q and not by P would inevitably come into existence ; and who (b) *did not believe* that that this choice would be wrong ; then it would have been, or in fact was, his *duty* to make that choice.

Fourth Pair. (i) The world is *intrinsically better* because A chose to do Y, when he could have chosen to do something else instead, than it would have been if he had made any other choice open to him. (ii) A *did his duty* in choosing Y.

Moore holds that in each of these four pairs the two members are interconnected by *synthetic* mutual entailment, but are not identical in meaning. If either the first or the second or the third were *analytic*, it would provide a *definition* of intrinsic value in terms of obligation. If the fourth were *analytic*, it would provide a definition of obligation in terms of intrinsic value. For the reasons given, Moore does not regard any of them as analytic, and he therefore sees no reason to think that either notion can be defined in terms of the other.

(IV) *Ethical Egoism and Ethical Neutralism*

In my essay in the *G. E. Moore* volume I defined what I call 'Ethical Neutralism' and what I call 'Ethical Egoism'. I pointed out that the latter might take milder or more extreme forms, but that it is in all its forms incompatible with Ethical Neutralism. I thought that Moore had claimed in *Principia Ethica* (pp. 96 to 105) to show that Ethical Egoism (at any rate in its extreme form) is *self-contradictory*. I argued that his attempt was a failure, and that all that could be proved was the tame proposition that Ethical Egoism is inconsistent with Ethical Neutralism.

Moore says, in his terminal essay, that what he was really trying to prove was not that Ethical Egoism *is self-contradictory*; but that Ethical Neutralism *would entail* that Ethical Egoism is self-contradictory, and not merely that it is false. Now Ethical Neutralism is at any rate highly plausible, and to some eminent moralists it has seemed self-evident on inspection. Therefore, if this argument of Moore's were acceptable, it would be at least highly plausible to hold that Ethical Egoism *is* self-contradictory, and not merely false.

Moore admits that his argument in *Principia Ethica* is extremely obscure and confused. He now produces a new argument and it is with this that we shall be concerned. It is extremely complex and hard to follow, and I am inclined to suspect that it contains a logical fallacy. In order to try to show this as clearly as possible, I shall exhibit formally what I take to be Moore's new argument.

In what follows I shall write 'p *ent* q' for 'p entails q', and I shall understand by this a kind of logical relation which holds, *e.g.* between the conjunction of the premisses of a valid syllogism and its conclusion. One way of describing it would be to say that the *conjunction* of p with not-q would be *impossible*, and that this impossibility does not depend on p being itself impossible

or on q being itself necessary. I shall write ' $p \text{ imp } q$ ' for ' p implies q ', and I shall understand by this that the conjunction of p with not- q is *in fact false*. With these notational preliminaries, the argument may be stated as follows.

Let ' p ' stand for the sentence 'It would *not be wrong* for X to choose Y '.

Let ' q ' stand for the sentence ' X does *not know* that the world would be *intrinsically worse* if he were to choose Y than if he were to choose some other alternative open to him at the time'.

Let ' r ' stand for the sentence ' X knows that the choice of Y by him would procure *for himself* a more favourable balance of intrinsically good over intrinsically bad experiences than any other choice that he could make, and knows also that this choice would be at least as favourable to the development of *his own* nature and dispositions as any other that he could make'.

Then what Moore calls 'Proposition A' is that p would follow from r *alone*, even though q should be false. So we may write ' A ' for ' $r \text{ ent } p$, even though not- q '. Moore asserts, and I agree, that A is entailed by Ethical Egoism.

What Moore calls 'Proposition B' is that the falsity of q would entail the falsity of p . So we may write ' B ' for 'not- $q \text{ ent } \text{not-}p$ '. Moore asserts, and I agree, that B is entailed by Ethical Neutralism.

Now Moore asserts that A entails that r does *not* entail q . For, he argues, to say that r would entail p , *even though q were false*, entails that it is *logically possible* for q to be *false*, even though r were *true*. This contention of Moore's may be written

$$A \text{ ent not-}(r \text{ ent } q)$$

Moore's argument may now be stated formally as follows :—

$$B \text{ ent (not-}q \text{ ent not-}p) \quad (\text{by definition})$$

$$\text{Hence } B \text{ ent } (p \text{ ent } q) \quad (\text{by contraposition}) \quad (\text{I})$$

$$\text{Again, } A \text{ ent } (r \text{ ent } p) \quad (\text{by definition})$$

$$\text{Therefore } (A \ \& \ B) \text{ ent } ((r \text{ ent } p) \ \& \ (p \text{ ent } q))$$

$$\text{Whence } (A \ \& \ B) \text{ ent } (r \text{ ent } q) \quad (\text{II})$$

But, as we have seen, according to Moore

$$A \text{ ent not-}(r \text{ ent } q)$$

$$\text{Therefore } (r \text{ ent } q) \text{ ent not-}A \quad (\text{by contraposition}) \quad (\text{III})$$

$$\text{Therefore } (\text{II}) \ \& \ (\text{III}) \text{ ent } ((A \ \& \ B) \text{ ent not-}A)$$

Since (II) and (III) can be asserted, we can drop them and assert what they together entail, *i.e.*

$$(A \ \& \ B) \text{ ent not-}A \quad (\text{IV})$$

Now up to this point the argument is valid, if we grant Moore's contention (which I shall not here question) that $A \text{ ent not-}(r \text{ ent } q)$.

But what he claims to have proved is that B entails that A is *self-contradictory*. Now this must be the proposition

B *ent* (A *ent* not-A) (V)

(It must be clearly understood that it is *not* enough for Moore to show that

B *ent* (A *imp* not-A)

For A *imp* not-A is simply equivalent to (not-A or not-A), which is in turn simply equivalent to not-A. So the latter proposition would merely amount to the tame conclusion that B *ent* not-A, *i.e.* in effect, that Ethical Neutralism is incompatible with Ethical Egoism.)

The question is, therefore, whether it is justifiable to infer from (IV), *i.e.* from (A & B) *ent* not-A, to (V), *i.e.* to B *ent* (A *ent* not-A). The answer is that this is *not* justifiable. Consider, *e.g.* a valid syllogism (P & Q) *ent* R. Suppose that you could legitimately derive from this the proposition P *ent* (Q *ent* R). Then, if the premiss P were known to be true, you could drop it and assert the proposition Q *ent* R. That this is not justifiable can easily be seen by taking a concrete example of a valid syllogism with a premiss known to be true. Take, *e.g.* 'All men are mortal' for 'P', 'Socrates is a man' for 'Q', and 'Socrates is mortal' for 'R'. Then, if this kind of inference were valid, we could infer from the syllogism the proposition (All men are mortal) *ent* ((Socrates is a man) *ent* (Socrates is mortal)). Then, dropping the true premiss that all men are mortal, we could assert that Socrates is a man *entails* that Socrates is mortal. Now that conclusion is certainly false. The mortality of Socrates is not a *necessary consequence* of his humanity *alone*.

So, unless I am much mistaken, Moore's new argument is fallacious, and he has failed to show that, if Ethical Neutralism were *true*, Ethical Egoism would be *self-contradictory*. It is a rash undertaking to accuse Moore of a logical fallacy, and it may well be that I have misunderstood his argument. On the other hand, it is very easy for the best of us to commit fallacies in *modal* logic, and so even Moore may have done so. But that consideration cuts both ways, and I myself may have committed some fallacy in modal logic in my criticism of his argument.

University of Cambridge