

## ***The Open Question Argument:***

### ***What it Isn't; and What it Is<sup>1</sup>***

#### 1. Introductory Comments.

Every competent history of 20<sup>th</sup> Century moral philosophy begins with a discussion of G. E. Moore's Open Question Argument (or "OQA"). According to a standard view, this argument set the stage for the next hundred years of philosophical reflection on the meanings of the terms of moral appraisal. Every metaethical theory that emerged in this period can be seen as some sort of reaction to OQA. Under the circumstances, one might assume that leading professional moral philosophers would be in fundamental agreement about how the argument is supposed to work, even if they do not entirely agree on its cogency.

Unfortunately, this assumption would be false. In their discussions of OQA, even the most respected of philosophers often attribute utterly unMoorean arguments to Moore. One particularly troubling pattern is common: a philosopher purports to be discussing OQA; he presents a confused argument unlike anything to be found in *Principia Ethica*; he then dismisses the argument with something approaching contempt, suggesting that only a simpleton could have been impressed by this "Moorean" argument. Surely, if one is going to shower contempt on Moore's argument, one might at least do him the courtesy of showering contempt on an argument he actually presented.

In this paper I attempt to take a small step toward rectifying this unfortunate situation.<sup>2</sup> I start in Section 2 with a brief review of the relevant passages in *Principia Ethica*. I then (in Section 3) make a few remarks on terminology. Then in Sections 4 and 5 I present two typical dismissive interpretations of Moore's argument. I point out the differences between what is actually in the text and what is

alleged by these interpretations. Then in Section 6 I present and discuss what I take to be the correct interpretations of the arguments. (I say ‘arguments’ because, as I see it, Moore actually presented several interestingly different versions of the argument.) In Section 7 I present what I take to be the “core” OQA. I conclude with a brief remark about the significance or importance of OQA.

## 2. Brief review of relevant texts.

If we are interested in Moore’s OQA, we must focus on Section 13 of *Principia Ethica*.<sup>3</sup> That Section contains an introductory paragraph followed by two subsections, numbered (1) and (2). Let’s first look at subsection (1). In this passage, Moore is talking about metaethical theories according to which goodness is “a complex whole”, something “definable”. Following Frankena, we can describe such theories as forms of “definism”. Moore’s stalking horse is (as he describes it) “the theory that to be good may mean to be that which we desire to desire”. Though he does not call it so, we can call this theory ‘DD’. He says a bunch of very dark things about this, and just past the middle of the subsection he mentions two questions:

Q1: Is it good to desire to desire A?

Q2: Do we desire to desire to desire to desire A?

He seems to say that these are different questions, as is shown by the fact that one is more complicated than the other. This apparently shows that DD is false. Moore suggests that DD is just an “instance” of the general definist approach. Presumably, he means to suggest that his objections to DD would be equally effective against other forms of definism. The conclusion of the argument seems to be that it is a mistake to suppose that “disagreement about the meaning of good is disagreement with regard to the correct analysis of given whole”. Let’s call this passage “the DD Passage”.

Immediately he says ‘moreover...’ and he makes a number of remarks about the conclusion that the predicate ‘good’ is (as he says) “positively different from” the notion of ‘desiring to desire’. He mentions another pair of sentences – these sentences are not questions:

S1: That we should desire to desire A is good.

S2: That A should be good is good.

He says that these are “*not* merely equivalent” and suggests that this bears on the truth of DD. He also mentions the proposition that all and only good things are things we desire to desire. His argument in the final sentence of this subsection of Section 13 seems to turn on the claim that we understand very well what is meant by doubting this proposition. He says that this shows that when we think of the subject (‘good things’) we have before our minds a different notion than when we think of its predicate (‘things we desire to desire’). Let’s call this passage “the Moreover Passage”.

Then in subsection (2) Moore turns to a different view which he (strangely) calls “the hypothesis that ‘good’ has no meaning whatsoever.” Presumably to illustrate this hypothesis he introduces a new stalking horse – the pleasure theory – according to which the notion of pleasure is identical to the notion of goodness. Here he mentions two other questions:

Q3: Is pleasure pleasant?

Q4: Is pleasure good?

He claims that everyone who thinks about it will see that these are two different questions; that the state of mind of one asking Q3 is different from the state of mind of one asking Q4; that the objects are different. He promptly repeats the exercise with a new pair of questions:

Q5: Is this good?

Q6: Is this pleasant?

Again he insists that if you think about it you will see that these are different questions. He seems to conclude either that this shows that 'good' has some meaning (outset of paragraph) or that the concept of goodness is distinct from other concepts (end of paragraph) or that 'good' is indefinable (beginning of next paragraph). As before, he makes it clear that the specific theory under attack – here the theory that pleasure is identical to goodness – is just an example of a certain class of theories. Let us call this passage “the No-meaning Passage”.

The phrase 'open question' does not appear in any of these passages, but it does appear a page later at the end of Section 14. Insofar as *Principia Ethica* is concerned, if there is anything that deserves to be called 'the Open Question Argument' in the book, it is to be found in these passages.<sup>4</sup>

### 3. Four preliminary problems.

There are several preliminary problems that should be mentioned before we begin in earnest to reflect on what's going on in these passages.

(a) Moore seems to be oblivious to the use/mention distinction. In some places he seems to be talking about words, but when he mentions those words they are not in quotation marks.<sup>5</sup> In other places he seems to be talking about concepts or properties, but he puts quotation marks around the words.<sup>6</sup> As a result, it appears that he might be talking but the words. In yet other cases, he apparently

follows the conventional procedure, using quotation marks to indicate mention of a word and absence of quotation marks to indicate use. This is a possible source of misunderstanding. I propose to follow the standard rules for use and mention. I use single quotation marks for mention of a word, and no quotation marks for use of a word. I use double quotation marks as scare quotes.

(b) Moore seems to be oblivious to the distinction between the use of a term as an ordinary predicate, like the phrase 'is left-handed' in 'Bob is left-handed', and the use of a term as a name of the property expressed by the predicate, like the noun 'left-handedness' in 'Left-handedness can be an advantage for a pitcher'.<sup>7</sup> This can create confusion. If someone were to say 'Lefthanded is my favorite' you might not know whether he was struggling to say that lefthandedness is his favorite property, or on the other hand, that his favorite person was left-handed.

I try to abide by the Ness-Itty-Hood Principle. Suppose I want to talk about some property but I'm not already familiar with any name for that property. I find a sentence whose predicate expresses that property. So, for example, suppose I am interested in the property that is distinctive of left-handers. Then I could start with 'Bob is left-handed'. Then I could take the predicate expression ('left-handed'), add 'ness' or 'ity' or 'hood' as appropriate. Voila! I have the name 'left-handedness'.<sup>8</sup> There are alternatives. I can use 'the property of being left-handed'. I can also use 'to be left-handed'. Moore seems unaware of any such rules. He says such things as 'good is a simple notion'. Does he mean that goodness is a simple notion? Or that it would be good to have a simple notion? Surely the former. If so, Moore has jammed a predicate expression ('good') into a slot that calls for a noun ('goodness'). He needs to add a '-ness' or an '-ity' or a '-hood'.

(c) Moore writes in such a way as to suggest that he does not fully appreciate the difference between the 'is' of identity and the 'is' of predication.<sup>9</sup>

Confusion of the 'is' of identity with the 'is' of predication creates special problems when combined with the other mistakes. So, for example, suppose someone says 'pleasure is good'. You don't know if he means to be saying: pleasantness = goodness, or whether instead he means to be saying that pleasantness is (predicatively) good.<sup>10</sup> I will try to highlight the relevant use of 'is' when it is important.

(d) Moore seems to be inconsistent in his use of linguistic terminology. The passages under consideration contain many occurrences of such terms as 'denote', 'mean', 'proposition', 'notion', 'predicate', etc. These terms are used in a lot of different ways by different philosophers. Some explicitly adopt certain of them for special uses. Moore's use is hard to fathom. My choices do not indicate claims to correctness. It's just that we have to agree about how we are going to talk. Consider 'Bob is left-handed'. The name 'Bob' *refers* to Bob. It's not clear that the predicate 'is left-handed' refers to anything. Let's say it *expresses* left-handedness. If you want a term that refers to left-handedness, why not say that 'left-handedness' refers to left-handedness? We can also say that 'left-handedness' *names* left-handedness.

Now let us consider two dismissive interpretations of OQA.

#### 4. Sturgeon's Identity Interpretation.

A number of philosophers have endorsed an interpretation of OQA according to which it turns essentially on a claim about the openness of certain identity questions. Nicholas Sturgeon has given this sort of interpretation of the argument. He says:

The simpler to explain, and to dismiss, is Moore's famous and widely influential "open question" argument in *Principia Ethica*... Moore assumes (a) that each of the theories under attack will provide an

explicit reductive identification for moral properties, such as “Goodness = conformity to the nature of the universe” or “Goodness = pleasure”; (b) that such property-identifications can be true only if the terms flanking the identity sign are synonymous for a competent speaker; and (c) that the terms can be shown not to be synonymous by the mere fact that a competent speaker could conceivably doubt the identity statement – by the fact, that is, that it remains an “open question” whether the identification is correct.<sup>11</sup>

As Sturgeon characterizes it in this passage, then, OQA turns essentially on some alleged features of certain identity statements.<sup>12</sup> The statements in question would be true if the relevant metaethical theories were true; but the identity statements are shown not to be true by the mere fact that a competent speaker can doubt them.

Let us use the name ‘PH’ to indicate the metaethical theory that Goodness is pleasure.<sup>13</sup> Then one specific version of OQA under Sturgeon’s interpretation is this:

### *OQA - 1*

1. It is an open question whether ‘Pleasure = Goodness’ is true.
2. If it is an open question whether ‘Pleasure = Goodness’ is true, then  $\sim$ (Pleasure = Goodness).
3. If  $\sim$ (pleasure = goodness), then PH is false.
4. Therefore, PH is false.<sup>14</sup>

To say that it is an open question whether a statement is true, according to Sturgeon, apparently is to say that a competent speaker of the language, understanding the meanings of all the terms (fully

understanding the meaning of the statement) could still be in doubt about the truth of the statement. Such doubt is “imaginable”.

Sturgeon suggests that this is a terrible argument. He says it is ‘simple to dismiss’; that it shows nothing of interest. He says that ‘scientific research appears frequently to have discovered’ that there are true identity statements in which the terms are not synonymous.<sup>15</sup> Thus, in effect, the openness of ‘pleasure = Goodness’ does not entail that pleasure is not goodness. Although he has other objections to the argument, Sturgeon’s main objection is that line (2) is indefensible. In his discussion of OQA, William Lycan formulates the argument in a similar way and reaches a similar conclusion, though he expresses his contempt for the argument in even stronger terminology.<sup>16</sup>

The critics are right about the argument under this interpretation. There is no plausibility to premise (2). There are many familiar cases in which we understand an identity sentence but nevertheless remain in doubt about its truth value. This can happen even if we are acquainted with the referents of both terms. Consider the familiar cases of the Morning Star and the Evening Star, Samuel Clemens and Mark Twain; Hesperus and Phosphorus; Clark Kent and Superman; Bruce Wayne and Batman. This holds even if A and B name properties. (Consider, for example, ‘my favorite property = your favorite property’.) So Sturgeon, Lycan, and any others who have attributed this argument to Moore are clearly right about one thing: OQA–1 is ‘easy to dismiss’.

But a deeper question remains: is this a fair interpretation of Moore’s argument? I would say that it is clearly not. I have carefully reviewed the sentences in Section 13 of *Principia Ethica*. I find only one identity sentence in the section. That is the sentence ‘Pleasure is the good’, which appears about halfway through subsection (2). This sentence seems not to serve as an example or as an



element in a premise of OQA. Rather, Moore simply seems to mention it in passing, merely to affirm that it is false. I encourage skeptical readers to look at all the sample sentences in the passages – all the sentences about which Moore makes some claim of openness, or doubtability. Each is a *predication* of ‘x is good’ or ‘x is pleasant’ or ‘x is something we desire to desire’. So Sturgeon gets the subject matter of Moore’s argument wrong.

Furthermore, there is no evidence to support the attribution of (2) to Moore. He never suggests that when an identity sentence is open to doubt, it is false. I find no passage in which Moore even suggests that terms referring to the same object must be synonymous for a competent speaker.<sup>17</sup> One application of the Principle of Charity is sufficient to dismiss Sturgeon’s interpretation: Moore never says anything that commits him to (2); (2) is an instance of a naive and silly principle; therefore, we would do well to look for a more plausible interpretation.

##### 5. The Darwall, Gibbard and Railton Self-Predication Interpretation.

Stephen Darwall, Allan Gibbard, and Peter Railton (hereafter ‘DG&R’) contributed a major paper on 20<sup>th</sup> Century ethics to the centennial issue of *The Philosophical Review*.<sup>18</sup> In that paper, DG&R correctly place OQA at the heart of modern metaethics. On their interpretation the argument has nothing to do with the openness of *identity* statements. As they see it, the argument turns on whether a certain sort of *predicative* question is intelligible. Here is the passage in which they give their summary of the argument:

One asks of any purported account identifying some descriptive property or state P as the meaning of ‘good’ whether on careful reflection we do not in fact find that we understand the question, “*Is P really good?*” If this question is intelligible—even, it seems, to those who hold that having or being P is a

good thing (perhaps the only good thing) and who are moved to give nonlinguistic reasons in defense of a positive answer to the question – then, absent some further story, P could hardly just be what we *mean* by ‘good’. (DG&R, 116)

Let us take pleasantness as our example of a descriptive property that might be taken as the meaning of ‘good’. Then the general drift of the argument under the DG&R interpretation seems to be this: if the meaning of ‘good’ were pleasantness, then the question whether pleasantness is good would not be intelligible -- we would not understand the question. Presumably this question would be unintelligible because it would be a matter of self-predication, and such questions are never intelligible. But the question whether pleasantness is good is intelligible. We do understand what it means. So the meaning of ‘good’ is not pleasantness.

So the argument would then look something like this:

### ***OQA - 2***

1. The question whether pleasantness is good is intelligible.
2. If the question whether pleasantness is good is intelligible, then ‘good’ does not mean pleasantness.
3. If ‘good’ does not mean pleasantness, then PH is false.
4. Therefore, PH is false.

Aside from its conclusion, this argument has little in common with OQA-1. Where OQA-1 focused on the openness of alleged identity statements, this argument focuses on the intelligibility of

alleged self-predication questions. The operative principle here seems to be that if some property, F-ness, is the meaning of some term, G, then the question ‘Is F-ness (predicatively) G?’ is not intelligible.<sup>19</sup>

DG&R present a generalized version of the argument.<sup>20</sup> In the interest of time and space conservation, I will not discuss it here.

While DG&R recognize that Moore might have been onto something important, they say that his deployment of the argument is “accident-prone”. They say it relies on “antiquated views in semantics”. They quote – apparently without dissent -- a distinguished philosopher who says it is “invalid”.<sup>21</sup>

It seems pretty clear that the argument thus formulated is “accident-prone” (if by this we mean that it is of a form whose instances could lead to a false conclusion). The problem concerns the underlying assumption that where some property, F-ness, is the meaning of some term N, the question whether F-ness is (predicatively) N is not intelligible. In other words, the problem concerns the general principle about self-predication.

In some cases it is obvious that a property does not have itself. The property of *being a physical object* is not itself a physical object; it does not have the property of *being a physical object*. In other cases it certainly seems obvious that a property has itself: the property of *being an abstract object* is itself an abstract object; it has the property of *being an abstract object*. In some cases it is not clear what we should say: consider the property of *being easy to analyze*. Is it easy to analyze? Does it have the property of *being easy to analyze*? The answer is not clear, but the question is intelligible. And in virtually every case,

the question whether a property has itself is intelligible. At least it makes sense, even if the answer is obvious.

The implication is that from the fact that a proposed definition has an intelligible associated question, nothing follows about whether the proposed definition is true. Whether pleasantness is the meaning of 'good' or not, it still might make sense to ask whether pleasantness is good. So premise (2) of OQA-2 is seriously problematic. More exactly, it presupposes a general principle about self-predication that is unmotivated. From the fact that the question 'Is pleasantness good?' is intelligible, nothing follows about whether pleasantness is the meaning of the word 'good'.

But there is a deeper question: is this a plausible interpretation of any of Moore's arguments? My answer is that it is not. There are several reasons for this. One is that in at least two of the passages cited, Moore compares two questions (or statements) and argues that they are non-synonymous. From this he seems to infer that two terms that appear in the questions (or statements) are non-synonymous. I see nothing of this "double question" procedure in the DG&R interpretation. For another, we can see that in two of the passages Moore talks about something other than whether a question is intelligible. In DD the argument turns on degrees of complexity. One question is alleged to be more complex than another. In the Moreover Passage it is alleged that you can just see by inspection that two statements involve different concepts. And in the No Meaning Passage the claim is that your state of mind is different when you ask one question from what it is when you ask the other. None of these passages contains an argument that turns essentially on the claim that some self-predication question is intelligible.

I recognize that there is a short passage near the bottom of p. 16 where one remark suggests something about self-predication. Moore mentions the question whether pleasure is after all good. If PH were true, this would be a case of self-predication. But notice that Moore does not rest his case on the claim that the self-predication would be unintelligible if the definition were true. He never says that the question is unintelligible. What he says in this passage is that the question 'is pleasure pleasant?' is distinct in meaning from the question 'is pleasure good?' The irrelevance of self-predication is reinforced at the top of the next page where the item under consideration is no longer pleasure, but some unidentified "this". The sample questions become 'Is this pleasant?' and 'Is this good?' There is some hint of self-predication in the earlier pair of sentences, but the self-predication disappears in the second pair of sentences, while the main point apparently remains intact. This strongly suggests that self-predication is a red herring in this process.

My impression is that DG&R have misunderstood the structure of Moore's argument. They have focused on a question that Moore mentions -- 'is pleasure good?'. They thought that Moore was arguing from the intelligibility of this question to the conclusion that 'pleasant' does not mean the same as 'good'. But in fact Moore makes no claim about the intelligibility of the question in this passage. Rather, he claims that the cited question differs in meaning from the question whether pleasure is pleasant. Further premises are then needed to get to the conclusion that 'pleasant' does not mean the same as 'good'. And in any case this is only one version of the argument. Self-predication plays no role in the other versions.

6. What in fact is OQA?

Let us now take a closer look at the passages from *Principia Ethica*. My claim is that these passages contain several closely related versions of OQA. None of them employs any premise concerning the openness or intelligibility of any identity question. Although some of Moore's examples tangentially involve questions about self-predication, none of the arguments essentially involves any controversial claims about self-predication in general.

The passages are alike in certain respects. Each is based on a sample theory; each purports to reveal something that is alleged to be true of a wider class of defnivist (subsection (1)) or non-defnivist naturalistic (subsection (2)) theories about the meaning of 'x is good'. Every version of the argument makes essential use of a principle of compositionality for meaning – a principle to the effect that the meaning of a whole statement is determined by the meanings of the parts.<sup>22</sup> In most cases, Moore provides two questions or statements somehow related to the sample theory. In these cases, he proceeds to claim that the questions or statements are distinct in meaning, and then concludes (by implicit appeal to the principle of compositionality) that this shows that the allegedly synonymous terms are in fact not synonymous. Since the theory implies that they are synonymous, this shows that the theory is false. Each passage suggests a way of proceeding to all other proposed definitions or analyses.

Let us first look again at the passage I earlier identified as “the DD Passage”. The theory under consideration here is this:

DD: x is good =df. we desire to desire x.

Note that DD is not only a form of naturalism (identifying goodness with a psychological, and hence naturalistic, property) it is also a form of definism (identifying goodness with a complex, analyzable property). In the passage in question, Moore's point is to establish something about *definism* – the view that it is possible to give an analytical definition of 'x is good' in its central moral use.

Moore starts by describing a procedure that involves self-predication – or at least something that would be self-predication if DD were true. That is, he asks of the “complex so defined” whether it is itself good. This would give us the question:

Q0: Is the property of being something we desire to desire good?

But Moore does not ask this question. Instead he asks this question:

Q1: Is it good to desire to desire A?

This is not precisely a self-predication question, but it is close. And then, after a bit of stalling, he says it is apparent that the meaning of this question cannot be correctly analysed into

Q2: Is the desire to desire A something that we desire to desire?

He says that when we ask Q1, we have not before our minds anything so complicated as the question 'Do we desire to desire to desire to desire A?'

Clearly, then, the argument involves something close to self-predication but only incidentally. The real point is that if DD were true, the two cited questions would have the same meaning. But we can see that they don't have the same meaning, since the meaning of one is not so complicated as the meaning of the other. So DD is not true. Thus, we may state the actual argument of this passage as follows:

### ***OQA - 3***

1. Q1 does not mean the same as Q2.
2. If Q1 does not mean the same as Q2, then 'x is good' does not mean the same as 'we desire to desire x'.
3. If 'x is good' does not mean the same as 'we desire to desire x', then DD is false.
4. Therefore, DD is false.

Moore defends (1) by claiming that Q2 has a more complicated meaning than Q1. One could debate this claim, saying that perhaps the added complexity is merely "verbal" – a misleading result of the fact that the surface structure of Q2 is somewhat confusing.

Although he does not explicitly state any such principle, I think Moore was presupposing a principle of compositionality. According to such a principle, the meaning of a linguistic whole is a function of the meanings of its parts and the pattern of their arrangement. This is a very plausible view. It goes a long way to explaining how we can understand the meaning of new sentences that we have never seen before. We are able to do this because we understand the meanings of the meaningful parts of the sentence, and we understand the structure of the sentence.



If the principle of compositionality is true, then the substitution of a synonymous expression for another expression already in a sentence should not change the meaning of a sentence. In other words, if you start with a sentence ‘ \_\_\_\_\_E1 \_\_\_\_\_ ’ and you replace E1 with a synonymous expression, E2, the resulting sentence ‘ \_\_\_\_\_E2\_\_\_\_\_ ’ should mean the same as the sentence with which you started. This implies that if two sentences are alike except that one contains a certain expression E1 where the other contains another expression E2, and the sentences do not mean the same, then the phrases E1 and E2 do not mean the same. The difference in meaning in the two phrases accounts for the difference in meaning of the two whole sentences. I believe that Moore implicitly appeals to this principle in the argument. He claims that the two cited questions do not have the same meaning. Since the questions are in fact alike except that one contains ‘good’ where the other contains ‘something that we desire to desire’, the principle allows us to infer that ‘good’ does not mean the same as ‘something that we desire to desire’.<sup>23</sup> As a result, DD is false. After all, DD purports to be a *definition*, it says that ‘x is good’ has the same meaning as ‘we desire to desire x’.

Moore claims that this line of argument can be used against any theory that purports to give an analytical definition of ‘x is good’ (“whatever definition be offered”). What is shown in OQA-3 about DD could be shown, he says, about any other proposed definition – that is, any definition according to which ‘x is good’ means the same as some complex expression. The target in subsection (1) is apparently definism, not naturalism. For even if the definition involved some evaluative terms, the procedure could still be employed.<sup>24</sup>

Patient inspection of the text will reveal (a) that Moore does indeed mention precisely those two questions in this passage; (b) that he does say that they differ in meaning; (c) that he does say that

this is sufficient to show that the sample theory is wrong, and (d) that he does claim that the same form of argument can be used with respect to the questions generated by other defnivist theories.

There are no property identity sentences anywhere in the passage. Although Moore starts out in such a way as to suggest that his argument will make use of claims about self-predication<sup>25</sup> this theme quickly disappears. When he finally states the argument it turns on features of questions Q1 and Q2. Neither of these questions involves self-predication. Moore's argument does not turn on the alleged intelligibility or openness of these questions. Rather, Moore's point is that the meaning of one of the questions is less complicated than the meaning of the other question. This provides his basis for saying that the two questions differ in meaning. And that (together with the principle of compositionality) enables him to infer that 'x is good' means something other than 'we desire to desire x'.

Let us now turn to the Moreover Passage, which is a continuation of the passage just discussed.

The sample theory under consideration in the Moreover Passage is still DD. Moore starts by (mis)stating his conclusion: the "predicate" 'good', he says, is positively different from the notion of 'desiring to desire'.<sup>26</sup> Evidently, his point is that 'x is good' does not mean the same as 'we desire to desire x'. These expressions would have the same meaning if DD were true. He then mentions two claims. They are:

C1: That we should desire to desire A is good.

C2: That A should be good is good.

He says that anyone can tell by inspection that the underlined terms do not mean the same. I take this to be a slightly misplaced statement of the intended conclusion. But what is the argument? I think the argument is compressed into one sentence. Moore says, speaking of C1 and C2, that the former 'is *not* merely equivalent to' the latter. I take him to be saying that the two sentences are not equivalent *in meaning*. If so, we have another version of OQA, very much in the pattern of OQA-3:

#### ***OQA - 4***

1. C1 does not mean the same as C2.
2. If C1 does not mean the same as C2, then 'A is good' does not mean the same as 'we desire to desire A'.
3. If 'A is good' does not mean the same as 'we desire to desire A', then DD is false.
4. Therefore, DD is false.

My comments on OQA-3 apply equally to OQA-4. OQA-4 does not involve any identity statements or self-predication statements; it turns essentially on a claim about sentence synonymy; it makes implicit use of a principle of compositionality for meaning; it is valid.

In the immediately following long and convoluted sentence in the Moreover Passage, Moore suggests a slightly different argument. He takes the two underlined phrases from C1 and C2 – phrases that would be synonymous if DD were true – and links them in a new sentence:

GDD: What we desire to desire is always also good.<sup>27</sup>

He claims that GDD is doubtful; he says that the mere fact that we understand what is meant by doubting it shows that when we contemplate the meanings of the two underlined phrases we have two different notions before our minds. Let us say that a statement is 'doubtable' just in case we understand very well what is meant by doubting it. I think the argument goes like this:

*OQA - 5*

1. GDD is doubtable.
2. If GDD is doubtable, then 'x is good' does not mean the same as 'we desire to desire x'.
3. If 'x is good' does not mean the same as 'we desire to desire x', then DD is false.
4. Therefore, DD is false.

As before, I assume that this is just an example of an argument pattern. The same line of reasoning could be used against any defintist theory.

OQA-5 is interestingly different from the other versions of OQA in the text. This is the only version of the argument that does not start with a pair of similar but allegedly nonsynonymous sentences and then appeal to compositionality to infer the nonsynonymy of two terms. However (as I will show below) it would be pretty easy to construct such an argument on the basis of the materials explicitly contained in this passage. In any case, OQA-5 is an interesting argument. Note (a) that there are no identity sentences anywhere in the argument; (b) that there are no self-predication sentences anywhere in the argument; (c) that the argument is valid. The argument as stated seems not to involve compositionality. I suspect that premise (2) relies on a general principle about synonymous terms, viz.

'If A and B are synonymous, then  $\langle(x)(Ax \rightarrow Bx)\rangle$  is not doubtable.' I am not entirely convinced that this principle is true.

Let us move now to the No Meaning Passage. In this passage Moore is considering another possible metaethical theory. This one is an instance not of definism, but of non-definist naturalism, for this one identifies goodness with pleasantness, which Moore takes to be a simple, natural property. He muddies the waters first by suggesting that he is talking about theories according to which 'good' has no meaning whatsoever. Surely, he meant to say that he is talking about naturalistic theories according to which 'good' has no *definition* whatsoever – theories according to which goodness is a simple and unanalyzable natural quality. He further muddies the waters by introducing talk of “the good”. As he made clear just a few pages earlier, “the good” is the unique thing that *has* goodness, not goodness itself.<sup>28</sup> He should have said that the theory under consideration here is this:

PH:  $x$  is good =df.  $x$  is pleasant

In any case, after a bit of introductory commentary, he says:

But whoever will attentively consider with himself what is actually before his mind when he asks the question 'Is pleasure (or whatever it may be) after all good?' can easily satisfy himself that he is not merely wondering whether pleasure is pleasant. ... Everyone does in fact understand the question 'Is this good?' When he thinks of it, his state of mind is different from what it would be, were he asked 'Is this pleasant, or desired, or approved?' It has a distinct meaning for him, even though he may not recognize in what respect it is distinct. (*PE*, pp. 16-17)

Thus, in this passage Moore at first explicitly mentions two questions:

Q3: Is pleasure pleasant?

Q4: Is pleasure good?

I recognize that there is something approaching self-predication in these questions. However, immediately after he mentions these questions he makes it clear that self-predication is irrelevant. He indicates that essentially the same point could just as easily be made with a different pair of questions, which he also explicitly states:

Q5: Is this pleasant?

Q6: Is this good?

He also indicates other pairs of questions:

Q7: Is this desired?

Q6: Is this good?

Q8: Is this approved?

Q6: Is this good?

In each case, Moore's point is the same. He says that the two compared questions "have distinct meanings". He gives several reasons for this: anyone who "attentively considers these questions with himself" recognizes that his state of mind is different when he asks Q4 from what it is when he asks

Q3; just look closely, you will “see” that a different idea is before your mind; etc. Moore’s point is that in each case we have a pair of nonsynonymous questions; but the questions differ only with respect to terms that would be synonymous if the relevant naturalist theory were true. So the argument (framed in terms of Q5 and Q6, though any other pair would do as well) goes like this:

***OQA - 6***

1. Q5 does not mean the same as Q6.
2. If Q5 does not mean the same as Q6, then ‘x is good’ does not mean the same as ‘x is pleasant’.
3. If ‘x is good’ does not mean the same as ‘x is pleasant’, then PH is false.
4. Therefore, PH is false.

Moore says that you can perform the same experiment with “each suggested definition”, but surely he does not mean ‘analytical *definition*’ since in this passage he means to be talking about non-definist naturalistic accounts, and these do not involve analytical definitions. They involve alleged mere synonymies. He should have said that you can perform the same experiment with any proposed non-definist naturalist account of the meaning of ‘x is good’.

As I see it, the crucial theoretical move here is again an appeal to compositionality. In each pair of questions, we have word-for-word similarity except for the word ‘good’ and the word that would be synonymous with it if the naturalistic theory were true. Again, Moore’s point is that since the questions are not synonymous, the terms with respect to which they differ are not synonymous. Hence, the “definitions” are to be rejected.

Notice that under this proposed interpretation (a) the argument does not involve any identity statements; (b) that questions about the emptiness or triviality or unintelligibility of self-predications are not relevant to this argument; (c) it involves an argument from the non-synonymy of two questions to the non-synonymy of two terms – terms that would be synonymous if the theory under consideration were true; (d) it allegedly can be extended to all other forms of non-definist naturalism; and (e) it is valid. It is not open to any of the obvious objections mentioned by Sturgeon and DG&R.

## 7. But where is the “Open Question Argument?”

One great disappointment here is that in all this discussion of the Open Question Argument we have seen almost nothing of the concept of *open questions*. It is a pity. It is also somewhat confusing. In order to set this straight, I will present an argument that (so far as I know) Moore never gave. This one really deserves to be called “the Open Question Argument”.

Let us say that a question is “open” if it is possible for a reasonable person to fully understand the question and every word in it, but still to be in doubt about the correct answer. Let us consider these two questions:

Q9: Is it the case that every pleasant thing is pleasant?

Q10: Is it the case that every pleasant thing is good?

A Moorean could point out that Q9 is not an open question, but Q10 is open. This shows (he might claim) that Q9 and Q10 do not have the same meaning. One application of the principle of compositionality yields the result that ‘pleasant’ does not mean the same as ‘good’. And that refutes PH. So the argument goes like this:



### *OQA – 7*

1. Q10 is an open question but Q9 is not.
2. If Q10 is an open question but Q9 is not, then Q10 does not mean the same as Q9.
3. If Q10 does not mean the same as Q9, then 'good' does not mean the same as 'pleasant'.
4. If 'good' does not mean the same as 'pleasant', then PH is false.
5. Therefore, PH is false.

The Moorean could claim that the argument can be generalized. Imagine any naturalist or defintist account of the meaning of 'good':

GN:  $x$  is good =df.  $x$  is N,

where 'N' is replaced by some suitable naturalistic or complex predicate expression.

Construct two question schemas:

Q11: Is it the case that everything that is N is N?

Q12: Is it the case that everything that is N is good?

Now it should be obvious that it is possible to construct a generalized form of the argument. (I will spare you the tedious details. Just follow the pattern of OQA-7, just stated.) My own view is that the generalized form of the argument would be pretty impressive, but of course the ultimate decision

would turn on the truth of the first premise. Would it be correct to say, in *every* instance, that the question of the form illustrated by Q12 is open? Perhaps we have to run through them one by one to find out. But in any case, this is something that truly deserves the name ‘Open Question Argument’. I furthermore think (a) it is not easy to dismiss; (b) it is not invalid; (c) it does not involve the naive blunder of supposing that names of identicals are synonymous; (d) it does not involve self-predication; (e) if it is “accident-prone” (whatever precisely that may mean), someone will have to tell what accidents it is prone to suffer.<sup>29</sup>

#### 8. What is the significance of OQA?

In recent discussions of OQA, commentators sometimes suggest that discoveries concerning rigid designators, natural kind terms, the causal theory of reference, etc. have shown us that OQA is little more than a quaint relic.<sup>30</sup> Saul Kripke and Hilary Putnam are often mentioned. Putnam himself seems to endorse this line of thinking. In one place, he starts out by saying that Kripke’s ideas have a “devastating” impact on Moore’s argument.<sup>31</sup> He then makes what I take to be a very surprising remark. He discusses a sentence that is relevantly like:

C3: This item is not good even though it is one that we desire to desire.

He suggests that OQA makes use of a premise to the effect that C3 is not self-contradictory. Putnam then says,

All that one can validly infer from the fact that [C3] is not self-contradictory is that ‘good’ is not synonymous with [‘we desire to desire’] (not synonymous with *P*, for any term *P* in the physicalistic version of the world).<sup>32</sup>

Putnam starts out by saying that Kripke's ideas have a devastating impact on Moore's argument. He then mentions in passing that the most Moore's argument establishes is that 'good' is not synonymous with 'we desire to desire'. But what is astonishing here is that *this is precisely the conclusion that Moore was concerned to establish!* In other words, Putnam has acknowledged that Moore's argument succeeds in proving the thing Moore wanted it to prove, but he nevertheless claims that the argument has been "devastated" because it does not prove some other thing.<sup>33</sup>

Perhaps the critical point is this: OQA may establish that 'good' is not exactly synonymous with any complex or naturalistic expression; but that is a conclusion of no interest. Who cares about synonymy?

But if this is the point, there is an answer. Moore cared about synonymy. He cared about it because he was worried about a certain line of argument. He was worried about arguments in which someone attempts to establish a substantive thesis in axiology by appeal to a premise about meaning. That this was Moore's interest can be seen clearly in this passage, which occurs just a few pages before the OQA passages we have been discussing:

"...if I am right, then nobody can foist upon us such an axiom as that 'Pleasure is the only good' or that 'The good is the desired' on the pretence that this is 'the very meaning of the word.'" (PE, p. 7)

Clearly, Moore was very concerned about arguments like this:

### ***First Synonymy Argument***

1. 'x is good' means the same as 'x is desired'.
2. If 'x is good' means the same as 'x is desired', then something is good if and only if it is desired.
3. If something is good if and only if it is desired, then preferentism is true.
4. Therefore, preferentism is true.<sup>34</sup>

and arguments like this:

### ***Second Synonymy Argument***

1. 'x is good' means the same as 'x is pleasant'.
2. If 'x is good' means the same as 'x is pleasant', then something is good if and only if it is pleasant.
3. If something is good if and only if it is pleasant, then hedonism is true.
4. Therefore, hedonism is true.

Moore wanted to deny premise (1) in each of these arguments and of any other argument of similar form. He wanted to insist that you can't derive substantive conclusions in axiology from claims about the synonymy of 'good' with some other expression (especially if the other expression was either naturalistic or complex). Therefore, he cared about the synonymy. That's why he presented OQA. Even some of the most ardent critics of OQA apparently are willing to admit that it succeeds at the job it was designed to perform. Perhaps they should not have been quite so dismissive.

---

## *Notes*

<sup>1</sup> I had the pleasure of discussing the topics of this paper on many occasions with Jean-Paul Vessel. Many of the ideas I here present can be found in Jean-Paul's dissertation, which I had the honor of directing. At this late date it is hard to determine whether I learned these things from Jean-Paul, or whether it was the other way around, or whether they just emerged in conversation. In any case, I am very grateful to Jean-Paul for his generous contribution to my understanding of OQA, even if I am not entirely clear about the precise boundaries of that contribution. I am also grateful to Jason Raibley, Owen McLeod, and Chris Heathwood for helpful criticism and suggestions.

<sup>2</sup> I do not mean to suggest that I am the first to take this step. Casimir Lewy and Stephen Ball have offered helpful and careful interpretive comments in important papers; they are cited in the references below. In Soames (2003) Scott Soames makes a number of interpretive points similar to the ones I make later in this paper. (In my defense, I may point out that I presented a version of these points well back in the Twentieth Century in Feldman (1978). I have no reason to suppose that Soames is familiar with that book.)

<sup>3</sup> Chris Heathwood raised an interesting question about this. Since the term 'open question' does not appear in Section 13, what reason is there for saying that this is the section that contains OQA? Perhaps I should say this: when subsequent philosophers spoke of OQA, they invariably mentioned lines of argument they found (or thought they found) in Section 13. I mean to be talking about those arguments, call them what you will.

<sup>4</sup> There is another passage in Section 74 that seems to contain another statement of OQA, but that is a muddy passage. I think it's appropriate to look for evidence concerning OQA in the passages I discuss even though the phrase 'open question' does not appear in them.

<sup>5</sup> As in the first sentence of subsection (1), where he speaks of '... disagreement about the meaning of good...' I think he means to be speaking of disagreement about the meaning of 'good'.

<sup>6</sup> As in the previous Section where he asks, 'Does that bind us to say that 'sweet' is exactly the same thing as 'yellow'...?' I think he means to be asking whether it binds us to say that sweetness = yellowness.

<sup>7</sup> As for example at the bottom of p. 14 where he says, '... there is no meaning in saying that pleasure is good, unless good is something different from pleasure.' Surely he means to say that there is no meaning in saying that pleasure is good unless *goodness* is something different from *pleasantness*.

---

<sup>8</sup> I am puzzled by the fact that we would never add '-ity' to 'left-handed' to get the name 'left-handedity'. Nor would we add 'hood' to get 'left-handedhood'. I leave it to the linguists to explain this familiar but puzzling phenomenon.

<sup>9</sup> Although of course from his other writings we know perfectly well that he understood the difference.

<sup>10</sup> The confusing use of 'pleasure' here opens the door to other possibilities. Maybe he is not talking about pleasantness at all.

<sup>11</sup> Sturgeon (1992): 799-800.

<sup>12</sup> In Sturgeon (1985), Sturgeon gives a different account of OQA. There he says that Moore claimed that we can always imagine someone to disagree with any "halfway interesting definition" of 'good'. Sturgeon claims that Moore thought that this would show that the definition is mistaken. According to Sturgeon, Moore went on to say that every naturalistic definition is in this way open to doubt, and so they are all mistaken.(p. 25) That seems to be Sturgeon's interpretation of OQA in Sturgeon (1985). Sturgeon goes on to claim that Moore's argument does not 'show anything of interest'.(p. 26) I certainly agree that the cited argument is hopeless. But the attribution to Moore seems to me indefensible.

<sup>13</sup> Here I am following Sturgeon's usage. I would prefer to formulate PH as the theory that 'x is good' means the same as 'x is pleasant', or perhaps as the theory that goodness = pleasantness.

<sup>14</sup> The argument can be generalized. Consider any name of a natural property, N. Suppose some naturalist offers a naturalistic definition of this form:

GN: Goodness = N

Then the identity statement formed by writing N followed by '=', followed by 'goodness' will be open. In other words, it will be an open question whether <N = goodness> is true. The names are not synonymous for a competent speaker (in Sturgeon's words). Therefore, no matter what N is, GN is false. Sturgeon says that Moore held that this openness shows that every identity claim of the illustrated form is false.

<sup>15</sup> p. 800.

<sup>16</sup> See Lycan (1988): 200-1, where Lycan says that Moore's argument 'simply fails; it is bankrupt.' I thank Jean-Paul Vessel for bringing this passage to my attention.

---

<sup>17</sup> I have not been able to find a completely satisfactory electronic copy of *Principia Ethica*. As a result, I have not been able to do an electronic search for the phrase ‘competent speaker’. However, I have spent a fair amount of time searching through my hard copy of the book. I have not found that phrase. I am certain that it does not appear in Section 13.

<sup>18</sup> Darwall, Gibbard, and Railton (1992): 115-189.

<sup>19</sup> There is some use/mention confusion in my statement of the question, but I hope the intent is sufficiently clear.

<sup>20</sup> DG&R (1992): 116.

<sup>21</sup> The distinguished philosopher is Gilbert Harman, whose quoted remark about invalidity appears in his (1977): 19-20.

There is a further complication: the argument that Harman dismisses as invalid is dramatically different from the argument presented by DG&R.

<sup>22</sup> With one exception, to be noted below.

<sup>23</sup> I acknowledge that for pure formal accuracy, it would be necessary to shift some words around in the sentences.

<sup>24</sup> This creates problems for Moore, since in other places he seems to give analytical definitions of ‘x is intrinsically good’. See, for example, the passage in which he says that the question whether something is intrinsically good is equivalent to the question whether it ought to exist for its own sake. This appears on p. viii of the Preface. Furthermore, in Moore (1912) he seems to say that ‘intrinsically good’ means the same as ‘would be good if it existed in complete isolation’ (p. 44) and also that ‘A is intrinsically good’ means the same as ‘if we had to choose between an action whose sole consequence would be A and an action that would lead to nothing at all, it would be our obligation to choose the former’ (p. 42). As I see it, if OQA is sound, these remarks about synonymy cannot be right.

<sup>25</sup> ‘... it may be always asked, with significance, of the complex so defined, whether it is itself good.’

<sup>26</sup> I assume that what Moore intended to say in this passage is that the meaning of ‘x is good’ is different from the meaning of ‘x is something we desire to desire’. I warned that Moore is not careful about the use/mention distinction!

<sup>27</sup> I take GDD to mean the same as: (x)(we desire to desire x --> x is good).

<sup>28</sup> Moore explicitly draws the distinction in Section 9. He says ‘good’ is an adjective; the good is the thing to which the adjective must always apply. In other words, the good is the thing that is always good.

<sup>29</sup> It is interesting to note that in Moore (1932) Moore seems to say (p. 127) that all of his arguments about the indefinability of ‘good’ in *Principia Ethica* are fallacious. It certainly seems that he there means to condemn every version of OQA. That would be unfortunate.

---

<sup>30</sup> A useful discussion of some of this literature can be found in Stephen Ball (1988). Ball explicitly (and insightfully) discusses arguments to be found in Harman and Putnam.

<sup>31</sup> Putnam (1981): 208.

<sup>32</sup> Putnam (1981): 208. Cited and discussed in Vessel (2003): 36. Putnam in fact discusses 'is good' and 'is conducive to maximizing total utility'.

<sup>33</sup> Sturgeon seems to make approximately the same point in Surgeon (1985): 26. He acknowledges that OQA might show that no naturalistic expression is a "precise cognitive synonym" for 'good'. But he says that this is no embarrassment for any naturalist. He suggests that this is a trivial and irrelevant point.

<sup>34</sup> Evidence that this was Moore's aim can be found in many places in *Principia Ethica*. A typical good example is the passage on p. 78 where he is talking about Mill's proof. He attributes to Mill the claim that 'because good means desired, therefore the desired is good.' Surely, Moore is thinking of an argument like the First Synonymy Argument. His discussions of Spencer's theory that the good is the more evolved follow a similar pattern.

### ***References***

Ambrose, Alice and Morris Lazerowitz. (1970) *G. E. Moore: Essays in Retrospect*, London: George Allen & Unwin, Ltd.,

Baldwin, Thomas. (1990) *G. E. Moore*, London: Routledge.

Ball, Steven. (1988) "Reductionism in Ethics and Science: A Contemporary Look at the Open Question Argument," *American Philosophical Quarterly* 25, 3: 197-213.

Becker, Lawrence and Charlotte Becker. (1992) *The Encyclopedia of Ethics*, New York: Garland Publishing.



---

Copp, David. (1992) "Metaethics," in Becker & Becker, *The Encyclopedia of Ethics*, 790-798.

Darwall, Stephen, Allan Gibbard and Peter Railton. (1992) "Toward *Fin de siecle* Ethics: Some Trends," *The Philosophical Review*, Vol. 101, No. 1, Philosophy in Review: Essays on Contemporary Philosophy: 115-189.

Feldman, Fred. (1978) *Introductory Ethics*, Englewood Cliffs: Prentice-Hall.

William Frankena. (1973) *Ethics*, (2<sup>nd</sup> ed.) Englewood Cliffs: Prentice-Hall.

\_\_\_\_\_. (1939) "The Naturalistic Fallacy," *Mind* XLVIII; reprinted in Klemke (1969), 30-43.

Hancock, Roger. (1969) "The Refutation of Naturalism in Moore and Hare," *The Journal of Philosophy* 57: 326-334. Reprinted in Klemke (1969), 44-52.

Harman, Gilbert. (1977) *The Nature of Morality*, New York: Oxford University Press.

Horgan, Terence, and Mark Timmons. (1992) "Troubles for New Wave Moral Semantics: The 'Open Question Argument' Revived," *Philosophical Papers* 21: 153-175.

Klemke, E. D. (1969) *Studies in the Philosophy of G. E. Moore*, Chicago: Quadrangle Books.

Kripke, Saul. (1980) *Naming and Necessity*, Cambridge: Cambridge University Press.

---

Lewy, Casimir. (1970) "G. E. Moore on the Naturalistic Fallacy," in Ambrose & Lazerowitz (1970): 292-303.

Lycan, William. (1988) *Judgment and Justification*, New York: Cambridge University Press.

Moore, G. E. (1903) *Principia Ethica*, Cambridge: Cambridge University Press. Reprinted many times. Paperback edition London, Cambridge University Press, 1962.

\_\_\_\_\_. (1912) *Ethics*, London: Oxford University Press.

\_\_\_\_\_. (1932) "Is Goodness a Quality?" *Aristotelian Society Proceedings*, Supp. Vol.: 116-131.

Putnam, Hilary. (1981) *Reason, Truth, and History*, Cambridge: Cambridge University Press.

Regan, Tom. (1992) "G. E. Moore," in Becker & Becker (1992): 821-2.

Rosati, Connie. (1995) "Naturalism, Normativity, and the Open Question Argument," *Nous* 29: 46-70.

Schilpp, Paul Arthur. (ed.) (1942) *The Philosophy of G. E. Moore*, Evanston and Chicago: Northwestern University Press.

Snare, Frank. (1975) "The Open Question as a Linguistic Test," *Ratio* 17: 122-9.

---

Soames, Scott. (2003) *Philosophical Analysis in the Twentieth Century*, Princeton and Oxford: Princeton University Press.

Sobel, Jordan Howard. *Good and Gold*, unpublished manuscript available online at [http://www.utoronto.ca/~sobel/Gd\\_Gld/](http://www.utoronto.ca/~sobel/Gd_Gld/)

Sturgeon, Nicholas. (1992) "Metaphysics and Epistemology," in Becker & Becker (1992): 798-804.

\_\_\_\_\_. (1985) "Gibbard on Moral Judgment and Norms," *Ethics* 96, 1: 22-33.

Vessel, Jean-Paul. (2003) *Open Questions and Consequentialist Conditionals: Central Puzzles in Moorean Moral Philosophy*, Doctoral Dissertation submitted at the University of Massachusetts at Amherst, February, 2003.