

TWO ARGUMENTS FOR EMOTIVISM AND A METHODOLOGICAL MORAL

CHARLES PIGDEN
Philosophy, U. of Otago
Dunedin 9016, New Zealand
CHARLES.PIGDEN@OTAGO.AC.NZ

In 1913 Russell gave up on the Moorean good. But since naturalism was not an option, that left two alternatives: the error theory and non-cognitivism. Despite a brief flirtation with the error theory Russell preferred the non-cognitivist option, developing a form of emotivism according to which to say that something is good is to express the desire that everyone should desire it. But why emotivism rather than the error theory? Because emotivism sorts better with Russell's Fundamental Principle that the "sentences we can understand must be composed of words with whose meaning we are acquainted." I construct an argument for emotivism featuring the Fundamental Principle that closely parallels Ayer's verificationist argument in *Language, Truth, and Logic*. I contend that Russell's argument, like Ayer's, is vulnerable to a Moorean critique. This suggests an important moral: *revisionist theories of meaning such as verificationism and the Fundamental Principle are prima facie false. Any modus ponens from such a principle to a surprising semantic conclusion (such as emotivism) is trumped by a Moorean modus tollens from the negation of the surprising semantics to the negation of the revisionist principle.*

I *Russell's Choice: Why Emotivism rather than the Error Theory?*

"When I was young," writes Russell, "I agreed with G. E. Moore in believing in the objectivity of good and evil. Santayana's criticism, in a book called *Winds of Doctrine* [which Russell read in 1913] caused me to abandon this view, though I have never been able to be

as bland and comfortable about it as he was.”¹ From that time onwards he subscribed to what he called “the subjectivity of value” by which he seems to have meant some kind of moral anti-realism. He thought that this doctrine might prove socially beneficial. It was not so much the truth that would make us free as the widely acknowledged *absence* of an objective moral truth that would make us free. “For my part, I should wish to see in the world less cruelty, persecution, punishment, and moral reprobation than exists at present; to this end, I believe that a recognition of the subjectivity of ethics might conduce.”² This hope was disappointed as the Bolsheviks, who subscribed in their metaethical moments to something like the subjectivity of value, were much given to “cruelty, persecution, punishment, and moral reprobation”, a fact of which Russell became uncomfortably aware when he visited Russia in 1920.³ But Russell was no pragmatist, and the fact that the subjectivity of value proved to be less beneficial than he had hoped did not lead him to abandon it as false. However, Russell’s explicit arguments for the “subjectivity of value” are objections to objectivism rather than arguments for a rival hypothesis. Moore’s theory is wrong since it presupposes non-existent, non-natural properties of goodness and badness.⁴ Despite a brief Humean wobble in the 1950s,⁵ Russell does not seem to have regarded naturalism—identifying moral properties with natural properties and thereby reducing moral facts to natural facts—as a viable option, presumably because he continued to subscribe to Moore’s Open Question Argument (henceforward the OQA).⁶ But if naturalism is not an option and if there are no non-natural facts or properties, that still leaves two metaethical alternatives: (a) some kind of non-cognitivism, or (b) an error theory. After a brief flirtation with the error theory in 1922,⁷ Russell’s dominant view was to be a form of emotivism, and hence of non-cognitivism, though he did not develop a fully articulated theory

¹ RUSSELL, “George Santayana” (1953), *PfM*, p. 91; *Papers* 11: 205.

² RUSSELL, “North Staffs’ Praise of War” (1916); *Papers* 13: 325; *RoE*, p. 117.

³ PIGDEN, “Emotivism, Error and the Meta-ethics of Bolshevism” (2014).

⁴ See PIGDEN, “Russell’s Moral Philosophy” (2014), §9 and my comments in *RoE*, pp. 17–19.

⁵ For Russell’s Humean wobble see *HSEP*, especially Chs. 1–9, *RoE*, §21, pp. 151–63, PIGDEN, “Russell’s Moral Philosophy” (2014), §10 and PERKINS, “Bertrand Russell’s Moral Philosophy” (2019), §2.5, pp. 348–51.

⁶ See RUSSELL’s two reviews of *Principia Ethica* at *Papers* 4: 568 and 571–2.

⁷ See “Is There an Absolute Good?” (1922), *RoE*, §17, and 58 in *Papers* 9.

until 1935 in Chapter 9 of his irreligious tract, *Religion and Science*. Statements about goodness and badness (the most basic kind of moral judgment for the mature Russell) are really in the optative mood. They express wishes or desires and are not, therefore, truth-apt. To say that *X* is good is to express the desire that everyone should desire it: “Would that everyone desired *X*!”⁸

But why did Russell prefer non-cognitivism to the error theory? In *Russell on Ethics* I suggested some reasons why emotivism might have been more psychologically appealing, and in a recent article Ray Perkins has followed suit with some further speculations along much the same lines.⁹ But in this paper I want to suggest something more philosophical, an *internal* rather than *external* explanation (so to speak) of why the emotivist research programme triumphed over the error theory in Bertrand Russell’s brain.¹⁰ Let me stress (channelling Lakatos) that this is a *rational reconstruction* of an argument which *might* have motivated Russell; an argument from premisses that Russell *seems* to have accepted to a conclusion that he *explicitly avowed*. It is *not* an argument that Russell himself explicitly develops in any extant text. Thus in so far as I am advocating a biographical thesis, rather than assembling an argument out of Russellian materials, it is this: *something* like the argument I suggest *must* have been at the back of Russell’s mind to justify his preference for emotivism over the error theory. Otherwise the choice would be under-motivated. My claim is that we can derive an argument *for* emotivism (and against the error theory) from Russell’s Fundamental Principle “that [the] sentences we can understand must be composed of words with whose meaning we are acquainted.”¹¹

Given the Fundamental Principle, Russell can make sense of non-existent properties but not of non-natural predicates. At least, he cannot make sense of predicates that are not definable in terms of things

⁸ See RUSSELL, “Science and Ethics”, in *Religion and Science* (1935) and in *RoE*, Ch. 19, and “Power and Moral Codes”, in *Power* (1938) and *RoE*, Ch. 27; PIGDEN, “Russell’s Moral Philosophy” (2014), §§6, 8–9; and PERKINS (2019), § 2.4, pp. 344–8.

⁹ *RoE*, pp. 189–93; PIGDEN, “Bertrand Russell: Moral Philosopher or Unphilosophical Moralism” (2003), p. 501, and “Emotivism, Error and the Metaethics of Bolshevism” (2014); PERKINS, “Bertrand Russell’s Moral Philosophy”, pp. 340–2.

¹⁰ See LAKATOS, *The Methodology of Scientific Research Programme*, Vol. 1 (1978), Ch. 2, for this distinction.

¹¹ RUSSELL, “Reply to Criticisms” (1944), in SCHILPP, *The Philosophy of Bertrand Russell*, p. 692; *Papers* II: 27.

with which we are acquainted. Thus on the assumption that we are not acquainted with goodness (which we obviously cannot be if there is really no such thing), and on the assumption that “good” cannot be defined in terms of the things with which we *are* acquainted (which seems pretty plausible if it is not equivalent to any naturalistic predicate), then we cannot even *understand* the predicate “good”. Thus the Fundamental Principle excludes the error theory if combined with the OQA. Furthermore it is possible to construct an argument for emotivism featuring the Fundamental Principle that closely parallels Ayer’s verificationist argument for emotivism in *Language, Truth, and Logic*. In this paper I will be constructing just such an argument before going on to contend that the Russellian argument, like Ayer’s, is vulnerable to a Moorean critique. This suggests an important moral about revisionary theories of meaning and why there is a strong presumption that they are false.

2. *The Errors in Russell’s Error Theory*

But before going on to this compare-and-contrast exercise, I had better say more about Russell’s version of the error theory and what I take to be wrong with it. The problem, as I see it, is this. Given the OQA and the Fundamental Principle, Russell cannot readily define the property that he wants to deny. But given the Fundamental Principle, he cannot deny a property that he cannot define. What Russell *wants* to say is that there is no such thing as an absolute property of goodness *as conceived by Moore*, and consequently that goodness-claims are all false. But given his semantic theory, he can’t quite manage to say this. Moore, of course, was famous for the claim that “good” (at least in a certain range of uses) is indefinable. Now, if “good” really were indefinable, and if it had the status of a proper name (that is, of an ultimate referring expression), then in Russell’s book it could not be meaningful unless it had a meaning, namely the property to which it referred. And this would require that property to exist. Thus we could not say meaningfully (let alone truly) that there is no such property as goodness. Since Russell wants to say that there is in fact no such property as goodness, he is forced to conclude that “good” is not a name but an “incomplete symbol”, an expression that does not have to have a meaning or a referent to function meaningfully in the context of a sentence. Russell believes, in effect, that if an apparent name *might* be empty but can be used, whether empty or not,

to make meaningful true-or-false claims, then it is not really a name but a disguised definite description. Russell concedes to Moore that there is “no doubt that our ethical judgments claim objectivity”,¹² which presumably entails that “good” *can* be used to make meaningful true-or-false claims. And he wants to argue that it actually lacks a referent. Thus he is forced to the conclusion that “good” can’t be indefinable. For, in order to be a cognitively meaningful but non-referring expression, “good” must be a disguised definite description—in which case it is susceptible to analysis.

But what would such an analysis look like? In “Is There an Absolute Good?” Russell suggests that in each person’s idiolect “*M* is good” means that *M* has the property shared by *A*, *B*, *C*, ... (where *A*, *B*, *C*, etc. are the kinds of things that the speaker happens to approve of) and *not* by *X*, *Y*, *Z*, ... (where *X*, *Y*, *Z*, etc. are the kinds of things that the speaker *disapproves* of). As Russell is careful to stress, the emotions of approval and disapproval do not enter into the *meaning* of the proposition “*M* is good”, but only into its *genesis* (which is why I have bracketed the clauses about approval and disapproval in my statement of Russell’s theory). Thus “goodness” for each person means *the property shared by certain kinds of things* (the kinds of things he or she approves of), *and not by certain others* (the kinds of things he or she disapproves of) but the notions of approval and disapproval do not figure in the italicized definite description itself. According to Russell this entails that “when we define it as nearly as possible in accordance with the usage of absolutists, *all* propositions in which the word ‘good’ has a primary occurrence are false.”¹³ But there are several things wrong with this theory.

(1) Since different people approve and disapprove of different things, when I say that *M* is good and you reply that it isn’t, we may not be contradicting each other. For what I mean by “*M* is good” is that *M* has the property shared by *A*, *B*, *C*, etc. (the things that I approve of) and not by *X*, *Y*, *Z*, etc. (the things that I disapprove of); whereas what you mean is that *M* does *not* have the property shared by *D*, *E*, *F*, etc. (the things that you approve of) and not by *U*, *V*, *W*, etc. (the things that you disapprove of). But if I say that *M* is good and you say that *M* is *not* good, we don’t cease to contradict each other

¹² “Is There an Absolute Good?”, *RoE*, p. 123; *Papers* 9: 345.

¹³ *RoE*, p. 122; *Papers* 9: 345.

because we happen to approve (and disapprove) of different things. And a theory which suggests that we do flies in the face of the semantic facts. Perhaps we can get around this by taking the “we” in Russell’s analysis rather more seriously. He is not talking about how “goodness” is to be defined in each person’s idiolect, but about a collective definition to which we collectively subscribe. “Goodness” stands for the property (if any) shared by certain kinds of things (the things that we, as a community, approve of) and not by certain other kinds of things (the things that we, as a community, disapprove of). The problem with this suggestion, however, is that it would reduce moral dissidents (such as Russell himself) to incoherence. I could not coherently think or say that “*things which are A are not good*” if “good” means the property shared by *A*-things, *B*-things, etc. (these being the kinds of things of which my community approves). For that would amount to the claim that *A*-things do not have the property shared by *A*-things, *B*-things, etc., which is a contradiction in terms. If there is such a property, *A*-things undoubtedly have it. If, for example, my community collectively approves of women’s subordination, truth-telling and the rule of the priests, I could not coherently suppose that women’s subordination was not good—for that would amount to the view that women’s subordination did not have the property common to women’s subordination, truth-telling and the rule of the priests. But moral dissidence is not semantically incoherent, which means that this Russellian analysis has to be wrong.

(2) The second difficulty is that Russell’s analysis fails to do what it intended to do since so far from implying that good-judgments are all *false*, it suggests that many of them are trivially true. For if I approve of *A*, *B*, *C*, etc., then there are at least two properties that *A*, *B*, *C*, etc. share. The first is the property of *being either A or B or C or ...*, and the second is the property of *being what I approve of*. Thus there will be at least one property shared by all the things that I approve of (and *not* by the things that I disapprove of). So if I say that *M is good*, meaning by this that *M* has the property shared by *A*, *B*, *C*, etc. and not by *X*, *Y*, *Z*, etc., then “*M is good*” will be true if (a) it is the disjunctive property that I have in mind and (b) *M* has the property of *being either A or B or C*, etc. (where “*A*”, “*B*”, “*C*”, etc. are conceived as the names of kinds). And if I say that *M is good*, meaning by this that *M* has the property shared by *A*, *B*, *C*, etc. and not by *X*, *Y*, *Z*, etc., then “*M is good*” will be true if (a) it is the property of *being approved of by me*

that I have in mind and (b) my assertion is sincere; for if I really think that *M is good*, it follows automatically that I approve of it and hence that it possesses the property of *being approved of by me*. Perhaps Russell could get around this either by denying the existence of properties altogether or by adopting a very sparse ontology of properties that excluded both disjunctive properties such as *being A or B, or C*, etc. and psychological properties such as *being approved of by me*. But an error theory that says that there is no such property as goodness—either because there are no properties at all or hardly any properties—is not very interesting; moreover it does not seem to be the kind of error theory that Russell was trying to develop. (Plenty of predicates and property words refer; it's just that "good" and "goodness" don't make the grade.) Alternatively Russell could argue that judgments of the form "*M is good*" are false because they mean something like *M has the property shared by A, B, C, ...* but that since *A, B, C*, etc. share *multiple* properties there isn't *one* property which is *the* property that they all share. But clearly this is not what Russell is trying to say. What he is trying to say is that propositions of the form "*M is good*" are false because *there is no such property as goodness*, not because there are *too many* properties shared by the things we approve of to meet a supposed uniqueness constraint.

(3) To my mind these difficulties are symptomatic of a deeper problem. Russell is trying to give an empiricist analysis of non-empirical concept. "I have maintained a principle [says Russell], which still seems to me completely valid, to the effect that, if we can understand what a sentence means, it must be composed entirely of words denoting things with which we are acquainted or definable in terms of such words" (*MPD*, p. 169). But what are the things with which we are acquainted? On this topic Russell changed his mind, though the extent of the changes remains a moot point. In the 1910s he supposed that we were acquainted with sense-data, meaning by this not sensations but sensory stimuli, the immediate physiological precursors of sensation, sensations being what happens when embodied minds become *aware* of sense-data. (Sense-data are a subclass of sensibilia, which are something like the sense-data that we *would* have were our bodies in certain positions. Thus sense-data which are actual are a subset of sensibilia, most of which are merely possible.) After 1921, when Russell adopted neutral monism, the distinction between the physical sense-data and the mental sensations fell away, and the two

classes collapsed into one another, both to be subsequently replaced by “percepts”. (“Individual percepts are the basis of all our knowledge, and no method exists by which we can begin with data which are public to many observers”, *HK*, p. 22.) For present purposes these differences¹⁴ don’t make much difference. The key point is that, early and late, both in 1922 when he flirted with the error theory and in 1935 when he explicitly converted to emotivism, Russell continued to think that the things we are acquainted with are *appearances*, which are private to each individual, not physical objects as usually conceived. We are not acquainted with the dead star, since the dead star is dead and gone, but only with the *image* of the dead star; we are not acquainted with the lawnmower but only the sound that the lawnmower makes or perhaps just the auditory sensations *caused* by the (physical) sound that the lawnmower makes. When combined with the Fundamental Principle this means that for each individual, words are only meaningful if they can be defined in terms of the sensations, the memories and perhaps the universals that that individual can experience. This is very close to what Bennett calls the *meaning-empiricism* of Locke, Berkeley and Hume.¹⁵ Public language is only meaningful to any person if it can be translated back into that person’s *private* language, whose fundamental terms stand for sense-data, sensations or percepts plus any universals to which that person has intellectual access.

Now why is this a problem when it comes to defining “goodness” “as nearly as possible in accordance with the usage of absolutists”? Because Russell has to construct his definite description out of materials given in experience, and the concept he is trying to define is neither given in experience nor readily definable in terms of percepts or universals that *are* given in experience. Perkins tries to deal with this difficulty by developing an improved version of Russell’s analysis: goodness is “the absolute non-natural property called ‘good’ that we find in common in all and only our objects of approval” (Perkins, p. 341). Now there is an obvious problem with the second clause (“that we find in common in all and only our objects of approval”). Absolutists are not infallibilists. They are not committed to the view that the things that we actually approve of are good or that the things that we

¹⁴ Detailed in WAHL, “Sense-data, Sensibilia and Percepts” (2019).

¹⁵ BENNETT, *Locke, Berkeley, Hume* (1971), p. 26.

don't approve of are not. With any luck our sentiments of approbation *track* the non-natural property of goodness in a reasonably reliable manner, but most absolutists would concede that there are probably some things that we approve of that are bad and some things that we disapprove of that are good. (Indeed every absolutist is committed to the view that the people they disagree with either approve of the bad or disapprove of the good. That's pretty much what it is to get it wrong in ethics, and the people who disagree with them are obviously getting it wrong!) Thus for an absolutist, goodness cannot be defined in terms of the objects of our approval, because the objects of our approval are not necessarily good (nor are the objects of our disapproval necessarily bad). Perhaps Russell (or Perkins on his behalf) could get around this by defining goodness as the property possessed by all and only the *proper* objects of our approval. But either properness is a new evaluative notion which stands in equal need of analysis, or "the proper of objects of our approval" simply means the things which are really good, in which case our analysis contains the very concept we are trying to analyze. There is the further problem that the notion of approval presupposes the concept of goodness, since to approve of something is to think or feel that it is good or right. Thus the second clause of Perkins' revamped Russellian analysis—"that we find in common in all and only our objects of approval"—is fatally flawed since it implies a degree of ethical infallibility that most absolutists reject and is viciously circular into the bargain, since it presupposes the very concept of which it purports to be an analysis.

But it's the first clause of Perkins' definite description that illustrates the deep difficulty that I have been trying to develop: "the absolute non-natural property called 'good'". For how are "absolute" and "non-natural" to be defined in terms of the objects of our immediate experience? "Absolute" gestures toward mind-independence, so if minds can be defined in terms of our immediate experience, and if we can each help ourselves in the solitude of our private semantics to the notion of causality and the notion of negation, we *might* be able to define absoluteness in terms of causal independency (*not* being causally affected by thoughts, *not* being caused to exist by anything mental or physical, *not* being such that it can be caused to cease to be), but it is going to be a tough job. Much the same goes for "non-natural". It might, I suppose, be possible first to define a property of properties answering to our common concept of "natural" and then to define

non-natural as *not* having that property, but it is going to be difficult to do given the limited conceptual resources at Russell's disposal (or rather the limited conceptual resources that he thinks we *all* have at our individual disposals). The key point, however, is this. In constructing analyses of "good" and "goodness" as disguised definite descriptions "as nearly as possible in accordance with the usage of absolutists", Russell cannot simply help himself to the conceptual vocabulary of his absolutist opponents (as Perkins seems to think). Given the Fundamental Principle, he is only entitled to those concepts that can be reductively defined in terms of the items with which we are each acquainted (whatever exactly these are supposed to be). This is going to be hard work especially as he cannot employ the notions of approval and disapproval since they presuppose the very concepts of which he is trying to give a reductive analysis.

3. *Arguing for Emotivism: A. J. Ayer*

Despite its tone of iconoclastic modernism, Ayer's *Language, Truth, and Logic* (1936; 2nd ed., 1946) is a highly derivative work, and the chief argument for emotivism is largely derived from Moore. Ayer was an admirer of Moore from way back. He read *Principia Ethica* as a teenager, at the instigation of the Bloomsbury aesthete, Clive Bell, who advised his readers to "run out this very minute and order of copy of [Moore's] masterpiece".¹⁶ Ayer took this advice and, like the members of the Bloomsbury circle when the book was first published in 1903, he swallowed Moore whole. It was not until his second year at Oxford that he "began to doubt whether 'good' was a simple, indefinable non-natural quality".¹⁷ Ayer may have come to doubt Moore's conclusion, but he continued to accept a large part of Moore's OQA. Here is my formulation of that OQA:¹⁸

- (1) "Are *X* things good?" is a significant or open question for any naturalistic or metaphysical predicate "*X*" (whether simple or complex).
- (2) If two expressions (whether simple or complex) are synonymous, this is evident on reflection to every competent speaker.

¹⁶ Quoted in AYER, *Part of My Life* (1977), p. 54.

¹⁷ AYER, "My Mental Development" (1992), p. 11.

¹⁸ PIGDEN, "Desiring to Desire" (2007), pp. 250–1.

- (3) The meaning of a predicate or property word is the property for which it stands. Thus if two predicates or property words have distinct meanings, they name distinct properties.
From (1) and (2) it follows that
- (4) “Good” is not synonymous with any naturalistic or metaphysical predicate “*X*” (or “goodness” with any corresponding noun or noun-phrase “*X*-ness”).
From (3) and (4) it follows that
- (5) Goodness is not identical with any natural or metaphysical property of *X*-ness.

In effect, Ayer accepted premisses (1) and (2) and therefore sub-conclusion (4), extending the indefinability thesis from “good” to “ethical concepts” generally, including “right” and “wrong”. “We have already rejected the ‘naturalistic’ theories which are commonly supposed to provide the only alternative to ‘absolutism’ in ethics ... [and] begin by admitting that the fundamental ethical concepts are unanalyzable” (*Language, Truth, and Logic*, p. 107). In other words “good”, like the other moral concepts, is not synonymous with any naturalistic predicate “*X*”. But Ayer was a verificationist. Synthetic propositions—which is what moral judgments would appear to be—are supposed to be either verifiable or senseless. But if “good” is not synonymous with any naturalistic predicate “*X*”, then *a fortiori* it is not synonymous with any *empirical* predicate, by which I mean the kind of predicate that might contribute to the verification-conditions of an empirically verifiable proposition. Now if non-verifiable propositions are senseless—mere pseudo-propositions in fact—and if “good” cannot help determine the verification-conditions of a verifiable proposition, then it would appear to follow that “good” is factually senseless, too. And this is precisely Ayer’s thesis. Ethical concepts “are mere pseudo-concepts”. “The presence of an ethical symbol in a proposition adds nothing to its factual content.” But although ethical concepts are mere pseudo-concepts and though they contribute nothing to the *factual* content of the sentences in which they appear, they are not totally devoid of meaning. They have a meaning, but that meaning is non-descriptive. Their function is to express the speaker’s feelings of approval and disapproval, to arouse similar feelings in others and thus (indirectly) to stimulate action.

We can represent Ayer's argument as a supplement to the OQA, taking sub-conclusion (4) as premiss:

(4) "Good" is not synonymous with any naturalistic predicate "X".

(4.a) All empirical predicates are naturalistic predicates.

[Assumption—but a pretty safe bet!]

Therefore

(4.b) "Good" is not synonymous with any empirical predicate "X".

[From (4) and (4.a).]

(4.c) A predicate is factually meaningful if and only if it is synonymous with an empirical predicate.

[Assumption—a consequence of verificationism.]

Therefore

(4.d) "Good" is not factually meaningful. [From (4.c) and (4.b).]

(4.e) "Good" is not meaningless.

[Assumption: an obvious semantic fact.]

Therefore

(4.f) "Good" has a non-factual meaning. [From (4.d) and (4.e).]

This theory has the added merit, in Ayer's eyes, of *explaining* why it is that "good" is indefinable. The reason why (1) "Are X things good?" is an open question for any naturalistic "X", and the reason why (4) "good" is not synonymous with any naturalistic predicate is that "good" is not, properly speaking, a predicate at all, since it cannot contribute to the truth-conditions of a sentence. What we are dimly aware of when we recognize "Are X things good?" as open is that it is up to us whether to approve or disapprove of X. Facts are one thing, feelings are another, and there is no logical or analytic connection between the way things are and the way we feel about them.

4. *Arguing for Emotivism: Russell*

All this is old hat. But it is worth rehearsing Ayer's argument because Russell's emotivism may have been motivated by a similar line of thought. Russell, of course, was not a verificationist. In his opinion there are synthetic propositions which are both unverifiable and factually meaningful. His example is "It snowed on Manhattan Island on the first of January in the year 1 AD." Either it snowed that day or it didn't; both suppositions make perfect sense, though neither is likely

to be verified (see *IMT*₂, p. 277). But Russell believed in a semantic thesis, his Fundamental Principle, which could be combined with the claim that we are not acquainted with anything non-natural to produce a similar argument for emotivism. The Fundamental Principle in one formulation is “that [the] sentences we can understand must be composed of words with whose meaning we are acquainted.”¹⁹ This needs to be modified slightly if the argument is to work. We say not that the “*sentences* we can understand [when analyzed] must be composed of words with whose meaning we are acquainted”, but that the “*factually significant* [or truth-apt] sentences that we can understand must [ultimately] be composed of words with whose meanings we are acquainted.” If we combine this with the thesis that we are not acquainted with anything non-natural we have the beginnings of an argument *for* non-cognitivism, as opposed to an argument *against* objectivism. This too can be represented as a supplement to the OQA, taking sub-conclusion (4) as a premiss.

- (4) “Good” is not synonymous with any naturalistic predicate “X”.
- (4.g) All factually significant predicates are definable (in use) in terms of the sense-data or percepts and the universals (if any) with which we are acquainted.
[Assumption: This is a consequence of Russell’s Fundamental Principle, that to understand a proposition we must be acquainted with the referents of its ultimate constituents.]
- (4.h) We are not acquainted with anything non-natural / Everything we are acquainted with is natural.
[Assumption, motivated in part by the considerations developed in Russell’s critique of objectivism.]
- (4.i) If a predicate can be defined with reference to naturalistic entities, then it is a naturalistic predicate. [Assumption.]
Therefore
- (4.j) All factually significant predicates are naturalistic predicates.
[From (4.g), (4.h) and (4.i).]
Therefore
- (4.k) “Good” is not synonymous with any factually significant predicate.
[From (4) and (4.j).]

¹⁹ “Reply to Criticisms”, p. 692; *Papers* 11: 27.

(4.l) A predicate is factually significant if and only if it is synonymous with a factually significant predicate.

[Assumption, but uncontroversial.]

Therefore

(4.m) “Good” is not factually significant. [From (4.k) and (4.l).]

(4.n) “Good” has some kind of significance.

[Assumption: an obvious semantic fact.]

Therefore

(4.o) “Good” has a non-factual significance.

[From (4.m) and (4.n).]

Russell, like Ayer, could have reinforced his theory by *explaining* why “good” is indefinable. If “good” really is an optative operator, designed to express a certain kind of desire, then of course it cannot be given a naturalistic analysis. As for the Open Question, facts are one thing, desires are another, and there is no logical or analytic link between the way things are and the way I desire that everyone should desire them to be.

5. *How to Avoid Non-Natural Goodness*

Given that they agree with Moore’s argument down to sub-conclusion (4), how do Ayer and Russell avoid the conclusion that they both came to doubt, namely that “‘good’ was a simple, indefinable non-natural quality”? The inference from (4) to (5) in the OQA proceeds via premiss (3), and the obvious option for those wishing to avoid a non-natural property would be to reject that premiss. But Russell seems to have accepted it, at least in an amended form, and it is not obvious to me that Ayer would have rejected it either. Thus both would have accepted

(3′) The meaning of a predicate is the property for which it stands. Thus if two predicates or property words have distinct meanings they denote distinct properties.

But both would have insisted on an amendment along the following lines:

(3″) The meaning of a predicate is the property for which it stands, so long as that predicate is (a) a complete symbol and

(b) factually meaningful. Thus if two *complete and factually meaningful* predicates have distinct meanings they denote distinct properties.

Since both Ayer and the emotivist Russell²⁰ regarded “good” as non-factual, this blocks the inference from (4)—a non-natural predicate “good”—to (5)—a non-natural property of goodness.

6. *The Ayer/Russell Modus Ponens*

What are we to make of the Ayer/Russell arguments? Simplifying somewhat, we can represent them both as instances of *modus ponens*, combining a conditional with a conjunction of semantic and philosophical claims. Here is Ayer’s argument:

- (A.1) (a) “Good” is not synonymous with any naturalistic predicate “X” and (b) a predicate is factually significant if and only if it is synonymous with a naturalistic [= empirical] predicate.
- (A.2) *If* (a) “good” is not synonymous with any naturalistic predicate “X” and (b) a predicate is factually meaningful if and only if it is synonymous with a naturalistic [= empirical] predicate, *then* “good” is not factually significant, i.e. such as to figure in truth-apt sentences.
- (A.3) Therefore “good” is not factually significant, i.e. such as to figure in truth-apt sentences.

And here is Russell’s argument (or at least the argument that I have attributed to him):

- (R.1) (a) “Good” is not synonymous with any naturalistic predicate “X” and (b) a predicate is factually significant if and only if it is synonymous with a naturalistic predicate, that is, a predicate

²⁰ In effect, the error-theoretic Russell of 1922 blocked the inference from (4) to (5) by insisting on the amendment that the meaning of a predicate is only the property for which it stands if the predicate in question is a complete symbol rather than a disguised definite description, which is what he took “good” to be. “Without the theory of incomplete symbols, it seemed natural to infer, as Moore did, that, since propositions in which the word ‘good’ occurs have meaning, therefore the word ‘good’ has meaning; but this was a fallacy. And it is upon this fallacy, I think, that the most apparently cogent of Moore’s arguments rest.” See “Is There an Absolute Good?”, *RoE*, p. 123; *Papers* 9: 345.

definable (in use) in terms of the naturalistic objects and universals with which we are acquainted.

- (R.2) *If* (a) “good” is not synonymous with any naturalistic predicate “*X*” and (b) a predicate is factually meaningful if and only if it is synonymous with a naturalistic predicate, i.e. a predicate definable (in use) in terms of the naturalistic objects and universals with which we are acquainted, *then* “good” is not factually significant, i.e. such as to figure in truth-apt sentences.
- (R.3) Therefore “good” is not factually significant, i.e. such as to figure in truth-apt sentences.

In both cases, the first premiss combines Moore’s sub-conclusion (4), that “good” is not synonymous with any naturalistic predicate “*X*”, with a criterion of factual significance for predicates derived from a grand semantic theory—verificationism in the case of Ayer, and the Fundamental Principle in the case of Russell. In both cases the two conjuncts of the first premiss are highly controversial, not to say intellectually suspect. But in both cases the second conditional premiss is relatively uncontroversial: it simply affirms that if “good” is a non-natural predicate and if non-natural predicates are not factually significant, then “good” is not factually significant. Thus each argument is open to the kind of response that Moore made to the idealists and to sceptics such as Hume.

7. *The Moorean Modus Tollens*

The idealist argues that we don’t really have flesh-and-blood hands (since they are, in some sense, mental rather than material entities), and the sceptic argues that we cannot *know* that we have flesh-and-blood hands. In both cases the arguments proceed from highly controversial philosophical premisses. Moore waives his hands and proclaims that he does indeed have flesh-and-blood hands and, furthermore, that he *knows* that he does. Thus—on the assumption that their arguments are valid—at least some of the premisses to which his opponents appeal must be false. This is not mere dogmatism since Moore’s *modus tollens* has a lot more going for it than his opponents’ *modus ponens*. They argue that if philosophical principles $P_1 \dots P_n$ are true, Moore does not have flesh-and-blood hands or does not know that he does; but principles $P_1 \dots P_n$ are true; therefore Moore does

not have flesh-and-blood hands or does not know that he does. Moore argues that if philosophical principles $P_1 \dots P_n$ are true, he does not have flesh-and-blood hands or does not know that he does; but Moore *does* have flesh-and-blood hands and, furthermore, he knows that he does; therefore at least some of the principles $P_1 \dots P_n$ are false. Both sets of arguments may be valid, but Moore's is more likely to be sound, since his platitudes are a lot more plausible than the speculative philosophical principles to which his opponents appeal.²¹

Now a cognitivist about ethics—maybe even Moore himself—could reply to Ayer and Russell along much the same lines. (Let me stress that just as the Russellian argument *for* emotivism is not an argument that Russell explicitly develops, so the Moorean argument *against* emotivism is not an argument that Moore explicitly deploys. I am adapting an argument that Moore uses in *one* context and applying to another.) Here's the Moorean argument against Ayer:

- (A.1') "Good" is a factually significant predicate that can play a part in truth-apt sentences.
- (A.2) *If* (a) "good" is not synonymous with any naturalistic predicate "X", and (b) a predicate is factually meaningful if and only if it is synonymous with a naturalistic [= empirical] predicate, *then* "good" is not factually significant, i.e. such as to figure in truth-apt sentences.
- (A.3') Therefore *either* it is *not* the case that (a) "good" is not synonymous with any naturalistic predicate "X", *or* it is *not* the case that (b) a predicate is factually significant if and only if it is synonymous with a naturalistic [= empirical] predicate.

And here's the Moorean argument against Russell:

- (R.1') "Good" is a factually significant predicate that can play a part in truth-apt sentences.
- (R.2) *If* (a) "good" is not synonymous with any naturalistic predicate "X", and (b) a predicate is factually meaningful if and only if

²¹ See MOORE, *Some Main Problems of Philosophy* (1953), especially Ch. 6; MOORE, *Selected Writings* (1993), Essays 5, 7 and 9; LYCAN, "Moore Against the New Skeptics" (2001) and "Moore's Anti-Skeptical Strategies" (2007); and SOAMES, *Philosophical Analysis in the Twentieth Century*, Vol. 1 (2003), Chs. 1 and 2.

it is synonymous with a naturalistic predicate, i.e. a predicate definable (in use) in terms of the naturalistic objects and universals with which we are acquainted, *then* “good” is not factually significant, i.e. such as to figure in truth-apt sentences.

- (R.3') Therefore *either* it is *not* the case that (a) “good” is not synonymous with any naturalistic predicate “X”, *or* it is *not* the case that (b) a predicate is factually significant if and only if it is synonymous with a naturalistic predicate, i.e. a predicate definable (in use) in terms of the naturalistic objects and universals with which we are acquainted.

Both sets of argument are valid, but which is most likely to be sound? The prize must go to the Moorean *modus tollens* arguments rather than the *modus ponens* arguments of Ayer and Russell. Indeed, these two “Moorean” arguments are rather better than the original Moorean arguments on which they are modelled. The problem with Moore’s polemical strategy is that it would seem to preclude anyone’s ever arriving at surprising conclusions about anything. We can imagine a seventeenth-century Moore responding to Galileo as follows:

Master Galileo, you have proved beyond reasonable doubt that if your premisses are true then the earth moves. Your logic I allow is impeccable. But since it is clear as daylight that the earth does *not* move—since this is far more epistemically compelling than any of the premisses from which you argue to the contrary—the obvious conclusion must be that some of your premisses are false!

A contemporary of Galileo’s who argued along these lines would have insulated himself in error, as would a nineteenth-century biologist who tried to do down Darwin with similar arguments. W. G. Lycan, who both defends and deploys Moore’s polemical strategy, replies that there is a considerable difference between the argument of Galileo’s proto-Moorean opponent and the arguments of the actual Moore. Galileo’s arguments are derived from empirical premisses whereas the premisses of Moore’s idealist and sceptical opponents were (as Lycan put the point when defending Moore’s strategy in the course of a Q and A session) just “philosophy junk”. The first kind of premiss can prevail against common sense; not so, the second. The problem with this is that if you actually examine Galileo’s arguments, what looks

like “philosophy junk” is conspicuous by its presence. Galileo certainly appeals to premisses which transcend observation and even glories in reason’s ability to overcome the apparent teachings of experience. More generally, a polemical strategy that depends on considerations of plausibility is fallible at best, since a proposition can be plausible but false or implausible but true.

That said, it is *very* implausible to suppose that for most of history most of humankind have misunderstood our own words, taking “good” to be a factual predicate when in fact it was an optative operator or a device for expressing emotion. The world at large is independent of our intentions, which is why it is easy to get things wrong. But what we mean is in some sense up to us, which is why widespread and long-lasting error about the meanings of everyday words seems to be unlikely. “Good” has usually been regarded as a factually significant predicate that can play a part in truth-apt sentences. We could not coherently tell tales about a “tree of the knowledge of good and evil” (*Genesis* 2: 17) if statements about the good were semantically incapable of truth, and knowledge of good and evil a conceptual impossibility. (This is not to say that there *are* [non-trivial] truths about good and evil, let alone that we know them; only that neither moral truth nor knowledge is precluded by the meanings of the very words.) It is not just that premisses (A.1’) and (R.1’) are highly plausible, having the backing of common sense. In *this* case, it is difficult to tell a coherent story about how common sense could be wrong.²² Meaning explains use, and if “good” and “bad” are used *as if* they are cognitive predicates, they must have a meaning that makes this possible. And it is hard to see what that could be except a cognitive meaning. There does not seem to be much daylight between *seeming to be* a factual or cognitive predicate and *being* a factual or cognitive predicate.²³ Thus the initial premisses of the Moorean arguments are pretty solid, whilst the second premisses of the two arguments are agreed by both sides.

Contrast the arguments of Ayer and Russell. Premises (A.1) and (R.1) are both conjunctions in which one of the conjuncts is sub-conclusion (4) of the OQA. This is suspect in itself since it is derived from

²² See PIGDEN, “Desiring to Desire” (2007), §§6 and 7.

²³ See PIGDEN, “Identifying Goodness” (2012): where I make a similar point against Philippa Foot who insists that “good” isn’t really a predicative adjective even though it is widely used as if it were (“Utilitarianism and the Virtues” [1985]).

premiss (2)—that if two expressions are synonymous this is evident on reflection to every competent speaker. Premiss (2) is problematic because it leads straight to the Paradox of Analysis. The Paradox of Analysis is this. Given (2), it is impossible for a philosophical analysis to be both true and interesting. For if the analysis is *true*, then the *analysandum* (the expression to be analyzed) must be synonymous with the *analysans* (the analyzing phrase or sentence). In such a case, the analysis will tell us nothing new, since if two expressions are synonymous this is evident on reflection to every competent speaker. Suppose, on the other hand, that the analysis tells us something new and interesting; suppose, that is, that it is *not* evident to every competent speaker that the *analysandum* is synonymous with the *analysans*. Then by (2) the two expressions are *not* synonymous, since it is not evident to every competent speaker that they are. The usual response to this problem is to deny premiss (2), thus opening the way for non-obvious synonymies and hence for analyses that are both true and interesting.²⁴ Of course, (4) could be true and (2) false, but in the absence of the OQA we no longer have a compelling reason to believe it. It is hardly likely to trump our everyday intuition that “good” is a factual predicate.

What about the second conjuncts of the two arguments? In both (A.1) and (R.1), the second conjunct applies a speculative theory of meaning to the special case of predicates. For Ayer the theory is verificationism, the idea that a synthetic sentence is only factually meaningful if it is empirically verifiable. From this he derives the consequence that a *predicate* can only be factually meaningful if it has verifiable content, that is, if it can contribute to the verification-conditions of a verifiable proposition. Perhaps this inference is not as secure as Ayer takes it to be, but the real problem lies with the verificationism on which it depends.

8. *Verificationism and Its Discontents*

The idea that a synthetic proposition is only factually meaningful if it is verifiable follows from the Wittgensteinian slogan that “the sense of a proposition is the method of its verification”.²⁵ If a purported

²⁴ See PIGDEN, “Desiring to Desire”, for more Moore, the OQA and the Paradox of Analysis.

²⁵ MCGUINNESS, ed., *Ludwig Wittgenstein and the Vienna Circle* (1979), pp. 79, 227,

proposition *has* no method of verification, then it has no (factual) sense or meaning. But the Wittgensteinian slogan suffers from a devastating objection. Generally speaking, there is no such thing as *the* method of verification of a proposition; hence the method of its verification cannot constitute the proposition's meaning. Whether a given set of observations verifies a proposition depends upon what else we take for granted. On some assumptions, observations $O_1 \dots O_n$ will tend to confirm a proposition and on others not. As Quine famously put it, "our statements about the external world face the tribunal of experience, not individually, but only as a corporate body."²⁶ Other problems arose when philosophers tried to formulate the verificationist criterion of meaning precisely. Just what does it mean for a proposition to be verifiable (or falsifiable) by experience? Successive formulations either excluded what its inventors meant to include—such as scientific laws and findings—or included what they meant to exclude—such as obviously metaphysical pronouncements, including such absurdities as "The Absolute is lazy."²⁷

9. Ross's "Refutation" of Verificationism

But important as these difficulties of detail proved to be, a deeper criticism was suggested by Sir David Ross in his response to Ayer.²⁸ He argues that verificationism is self-refuting however formulated. For it claims that a sentence is only factually or cognitively meaningful if it is either analytic, contradictory or empirically verifiable (or perhaps falsifiable). But the verification principle itself—that a sentence is only factually or cognitively meaningful if it is either analytic, contradictory or empirically verifiable (or perhaps falsifiable)—is itself not analytic, contradictory or empirically verifiable or falsifiable. Thus if it is true, it is not cognitively meaningful, which means that it cannot be true.

Some verificationists tried to get out of this difficulty by suggesting that the verificationist criterion of meaning should be read as a proposal rather than a proposition. The claim is not that, as a matter of fact, non-verifiable sentences are factually or cognitively meaningless, but that it might be a *good idea* to *treat* them as if they were.

244, 247.

²⁶ QUINE, "Two Dogmas of Empiricism" (1951), p. 49.

²⁷ See SOAMES, pp. 277–91, for a potted history.

²⁸ ROSS, *The Foundations of Ethics* (1939), p. 38.

Verificationism is not a self-refuting attempt to tell it like it is (a purported truth which is factually meaningless if true, and so not true), but a non-cognitive suggestion about how some sentences should be regarded. But if verificationism degenerates into a mere suggestion, it loses its polemical bite since its opponents are at liberty to reject it. If you don't like Ayer's conclusions—and many don't—you can simply evade them by refusing to accept his proposal.²⁹

Cute as Ross's criticism is, I am inclined to think that it is mistaken. True, it is a clever instance of a nifty polemical strategy. Philosophers are wont to claim that kosher propositions are all of kind *K*, when the claim itself is not of the kosher kind *K*. And it is a generally a smart move to point this out. But not this time. For the verificationist criterion of factual meaning (broadly conceived) *does* meet its own standards for factual significance. (It claims that kosher—in this case *factually meaningful*—propositions are all of kind *K*, which *is* a kind of proposition to which it belongs.) Thus it is not (or need not be) self-refuting. True, it is not the kind of claim that can be conclusively verified or falsified (but then almost nothing is). But it *is* the kind of claim that can be confirmed or disconfirmed by the empirical evidence. Thus the problem is not that it is self-refuting. The problem is that it is empirically false.

10. *Why Revisionist Theories of Meaning Are Prima Facie False*

What is the task of a theory of meaning? To explain how it is that marks on paper or sound patterns in the air manage to mean what they mean. Thus the facts of everyday usage, and our everyday intuitions about what means what, constitute the data that a theory of

²⁹ Coffa quotes Reichenbach on the choice between two different conceptions of verifiability and hence (for him) of meaningfulness. "If we are now to make a choice between these two definitions ... we must clearly keep in our mind that this is a question for a volitional decision and not a question of truth-character. It would be entirely erroneous to ask What is the true conception of meaning? or which conception must I choose? Such questions would be meaningless [*sic!*] because meaning can only be determined by a definition. What we could do would be to propose the acceptance of this decision." Coffa goes on to observe: "This forgettable semantic doctrine was frequently forgotten by Reichenbach when he got down to the business of actual philosophizing, in which he was usually inspired by a rather extreme form of semantic intolerance. One often finds him arguing about the precise meaning of this or that claim with far greater ego involvement than the mere proposal of a convention could possibly inspire" (COFFA, *The Semantic Tradition from Kant to Carnap* [1991], pp. 317–18).

meaning has to explain. These are the facts that a theory of meaning has to fit. A theory of meaning for a language, *L*, would consist of two parts: *general* principles about the way words and phrases of various kinds work (how verbs manage to be meaningful, for instance) and *particular* theses relating to the vocabulary and grammar of *L*. The “predictions” (or retro-dictions) of such a theory should correspond to the contents of a good dictionary, which itself merely codifies the linguistic intuitions of educated speakers. A theory of meaning is confirmed (to some extent) if it matches those intuitions, and disconfirmed if it does not.

Now, the verification principle is a high-level principle about the way that language works, confining factual meaning to claims of certain kinds. It would be included in the general part of a wide range of theories of meaning for specific languages. *But its tendency would be to falsify such theories.* Why so? Because it is an explicitly *revisionist* thesis. A large part of the point of the verification principle is to exclude as meaningless many claims that are widely thought to be meaningful (such as “God exists”). It was designed by Wittgenstein and the Vienna positivists as a philosophical weapon of mass destruction which would allow them to dismiss the people they disagreed with without having to argue against them in detail. But this means that if it is incorporated into a larger theory of meaning, that theory will tend to issue in the false predictions. (Though which predictions it will issue in depends on which version of the principle we adopt.) It will “predict”, for example, that “It snowed on Manhattan Island on the first of January in the year 1 AD” lacks a truth-value, which evidently it does not. It will “predict” that (on certain assumptions) theism, atheism and agnosticism do not represent cognitively meaningful opinions (one of Ayer’s more startling conclusions), or that “it can not be significantly asserted [or significantly denied] that men have immortal souls” (Ayer, p. 31). It will “predict” (if we adopt Neurath’s variant) that terms like “cause”, “true” and “explanation” cannot be used to make cognitive claims. It would entail that we cannot give a rational reconstruction of large chunks of the past, since, in many cases, the thoughts which have directed people’s actions would be damned by the verificationist as lacking in cognitive content. You can’t explain someone’s actions as due to a belief that his soul is in danger, if it does not make sense to suppose that he has a soul. Indeed, if Ayer had been strict with himself, he would not have been able to make sense of his

own mental development since (as we have seen) it involved a non-sensical belief in a Moorean good. Thus history—or at least the kind of history which explains people’s actions in terms of their beliefs and desires—would be riddled with islands of unintelligibility and would be largely condemned as bunk: *not* a happy consequence. For verificationism implies that many of our utterances don’t mean anything at all and that others have a meaning that nobody dreamt of until the twentieth century.

Thus if verificationism is added to a theory of meaning, it will tend to issue in predictions that are false to the facts, namely our everyday intuitions about what words mean. It is as if we had a theory to explain how ducks fly which had the surprising consequence that many of them don’t, and that some that do fly, fly backwards. The reason for this empirical failure is not hard to find. Verificationism was never *designed* to explain the facts of linguistic usage but to modify those facts in the interests of a scientific agenda. It is as if we had a theory of how ducks fly whose covert purpose was to *stop* many of them from flying by persuading them that, for many ducks, flight is an impossibility. Paraphrasing Marx, we might say that verificationist philosophers have attempted to *change* linguistic usage in various ways, but the point of a theory of meaning is to *explain* it. Hence the empirical failure.³⁰ Thus the verification principle fails, partly because of Quine’s criticisms, but mainly because, in so far as it can be clarified, the theory is *not* self-refuting but empirically *false*. This is important for two reasons.

II. Ayer’s Argument Again

For a start the failure of the verification principle suggests that Ayer’s argument is unsound. Premiss (A.1) combines the Moorean thesis (a)—that “good” is not synonymous with any naturalistic predicate “X”—with the verificationist thesis (b)—that a predicate is factually significant if and only if it is synonymous with an empirical predicate, i.e. a predicate that can play a part in verifiable propositions. Thesis (a) is dubious since it is derived from premiss (2) of the OQA which leads straight to the Paradox of Analysis. And thesis (b) is false since it depends upon verificationism, which—in so far as it can

³⁰ See HENLE, “Meaning and Verifiability” (1963), for a similar argument; also PIGDEN, “Coercive Theories of Meaning” (2010).

be clarified—would appear to be empirically false. But there is more. In so far as verificationism leads to a surprising semantic conclusion such as emotivism, that is *evidence* that it is false. For a theory of meaning that aspires to fit the facts *should not* lead to such surprising semantic conclusions. A theory that “predicts” that moral judgments mean something that nobody had ever thought of till the advent of Russell and Ayer is a theory that is contradicted by the evidence, namely the evidence of our linguistic intuitions. This gives us a *principled* reason for preferring the Moorean *modus tollens* to Ayer’s *modus ponens*, and thus something better than flat-footed considerations of relative plausibility.

12. *Russell’s Argument and the Fundamental Principle*

What about Russell? The Russellian argument differs only slightly from Ayer’s. The premiss (R.1) combines thesis (a)—that “good” is not synonymous with any naturalistic predicate “*X*”—with thesis (b)—that a predicate is factually significant if and only if it is synonymous with a predicate definable (in use) in terms of the naturalistic objects and universals with which we are acquainted. Thesis (a) is dubious since it is derived from premiss (2) of the OQA which leads straight to the Paradox of Analysis. But what about thesis (b)? That depends on Russell’s Fundamental Principle “that [the] sentences we can understand must [ultimately] be composed of words with whose meaning we are acquainted”, together with the thesis that we are only acquainted with naturalistic things.

What can be said for the Fundamental Principle? Not much. It was the bane of Russell’s existence as a philosopher and was always getting him into trouble. Indeed, if you catch Russell saying something weird or implausible, the chances are that the Fundamental Principle is at the back of it. Take, for example, his widely lampooned doctrine that “this” is a proper name³¹. It is derived, in part, from the Fundamental Principle. (Proper names are the ultimate constituents of sentences: their meaning consists in the objects to which they refer. They are the words “which are only significant because there is something that they mean, and if there were not this something, they would be empty

³¹ Cf. *The Philosophical Lexicon*, 8th edition: “bertrand: a state of linguistic amnesia as of one who thinks ‘this’ is a proper name and ‘Plato’ a description.”

noises, not words.”³²) By the Fundamental Principle, we must be acquainted with these somethings. But in 1918, the only non-predicates that Russell could think of which referred directly to the objects of our acquaintance were words like “this” when used to denote sense-data. Hence “this”, in these uses, functions as a proper name.³³ Ditto, his equally bizarre doctrine that ordinary language is ambiguous, so that typically when “one person uses a word, he does not mean by it the same thing as another person means by it.”³⁴ Sometimes the Fundamental Principle drove Russell to the edge of despair:

But now all this [the universe] has shrunk to be no more than my own reflection in the windows of the soul through which I look out on the night of nothingness. The revolutions of the nebulae, the birth and death of stars, are no more than convenient fictions in the trivial work of linking together my own sensations, and perhaps those of other men not much better than myself. (*Auto.* 2: 158)

But it was not, as he thought, “the shadow physics of our time” that threatened to imprison Russell in this solipsistic “dungeon” (*ibid.*) but his own Fundamental Principle. For if the “sentences we can understand must be [ultimately] composed of words with whose meaning we are acquainted”, and if we are only acquainted with our own percepts, it is hard to see how we can talk or even *think* about stars or nebulae if these are conceived as mind-independent entities. In the end, Russell was able to escape this intellectual prison with the aid of his theory of definite descriptions (with external objects defined, in effect, as the causes of our percepts); but the point is that he would never have been at risk of arrest from the forces of solipsism had it not been for the Fundamental Principle.³⁵

Thus the Fundamental Principle is bad news, especially for a would-be scientific realist such as Russell. Furthermore, it does not sit well with Moore’s premiss (2), from which thesis (a)—that “good” is not synonymous with any naturalistic predicate “*X*”—is derived. If

³² RUSSELL, “Mr. Strawson on Referring” (1957), in *MPD*, pp. 240–1; *Papers* 11: 632.

³³ See *PLA*, in *LK*, pp. 200–2; *Papers* 8: 178–9.

³⁴ By the Fundamental Principle, “the meaning you attach to your words must depend on the nature of the objects you are acquainted with”, and “since different people are acquainted with different objects [or sense-data]” different people do not mean the same things by the same words. See *PLA*, in *LK*, p. 195; *Papers* 8: 174.

³⁵ See MAXWELL, “The Later Bertrand Russell” (1974), for further details.

the Fundamental Principle is correct, and if we are only acquainted with sense-data or percepts and universals, then “I am looking at page 158 of Russell’s *Autobiography*” is equivalent to a long and involved sentence about black and white percepts, in which the book itself features (if at all) as an unknown cause. Now it is certainly *not* obvious to every competent speaker that “I am looking at page 158 of Russell’s *Autobiography*” is equivalent to such a sentence. But by (2)—that if two expressions are synonymous this is evident on reflection to every competent speaker—if it is *not* evident to every competent speaker that two expressions are synonymous, it follows that they are not, in fact, synonymous. Hence “I am looking at page 158 of Russell’s *Autobiography*” is *not* equivalent to a long and involved sentence about black and white percepts. But if the two expressions are not equivalent, then Russell is wrong. For *either* the Fundamental Principle is false or he is wrong about acquaintance, which is not confined to percepts but can include such things as books.

This is not a decisive objection to the Russellian argument, however, since thesis (R.1)(a)—that “good” is not synonymous with any naturalistic predicate “*X*”—might be true and the thesis from which it is derived—(2) that if two expressions are synonymous this is evident on reflection to every competent speaker—might be false. Hence the fact that, on certain assumptions, (2) is incompatible with (R.1)(b) does not show that premiss (R.1), including theses (a) and (b), is inconsistent. But it does point to a fundamental problem with the Fundamental Principle.

13. *What’s Really Wrong with the Fundamental Principle*

The Fundamental Principle, like verificationism, is a high level thesis about how meaning works. It says that factually significant sentences are meaningful in virtue of being analyzable into sentences whose constituents refer directly to the objects of our acquaintance. These are taken to be sense-data or percepts, and perhaps the universals to which we have sensory or intellectual access. The Fundamental Principle would figure as one of the general principles shared by specific theories of meaning for a specific language *L*. Such theories are confirmed if they spit out predictions in broad agreement with our linguistic intuitions, and are disconfirmed if they do not. For it is the business of a theory of meaning to explain both our everyday intuitions about meaning and the facts of linguistic usage. But if this is the

case, the Fundamental Principle would tend to sabotage any theory to which it was added. For it issues in predictions that fly in the face of the linguistic facts. It is a consequence of the Fundamental Principle that “Charles owns a copy of Russell’s *Autobiography*” said by me means something very different from “Charles owns a copy of Russell’s *Autobiography*” said by you. Since our sentences are only factually meaningful because they can be cashed out in terms of the sense-data or percepts with which we are acquainted, and since we are each acquainted with different sense-data or percepts, our two utterances cannot possibly be equivalent. Yet our everyday intuition that the two sentences are equivalent is the kind of linguistic datum that a theory of meaning ought to predict. It is a consequence of the Fundamental Principle that “Charles owns a copy of Russell’s *Autobiography*” is equivalent to some long complex sentence about sense-data or percepts in which the book itself figures (if at all) as an unsensed cause, to be singled out by an elaborate definite description. Yet nobody has managed to formulate such a sentence, few would understand it if it *were* formulated, and even if *could* be formulated and *were* understood, many would be inclined to reject it as not what the original utterance was trying to say. (“I was not talking about the sense-data or percepts” they would insist, “I was talking about the *book* and the fact that Charles happens to own it.”) If a thesis about meaning conflicts with our linguistic intuitions, that is evidence that it is false, and the Fundamental Principle conflicts with our linguistic intuitions.

14. *The Russellian Argument Again*

This brings us back to the Russellian argument for non-cognitivism. This proceeds from two premisses, the uncontentious (R.2)—that if (a) “good” is not synonymous with any naturalistic predicate, and (b) a predicate is factually meaningful if and only if it definable (in use) in terms of the naturalistic objects and universals with which we are acquainted, *then* “good” is not factually significant—together with its much more contentious antecedent (R.1)—that (a) “good” is *not* synonymous with any naturalistic predicate, and (b) that a predicate is factually meaningful if and only if it definable (in use) in terms of the objects with which we are acquainted. As we have already noted, thesis (R.1)(a) is suspect since it is derived from premiss (2) of the OQA, which leads straight to the Paradox of Analysis. So too is thesis (R.1)(b) since it is derived from Russell’s Fundamental Principle, a

thesis about meaning which at odds with the empirical evidence. Thus the Russellian argument is, at best, highly dubious and, at worst, unsound, since one of the premisses appears to be false. But there is more. In so far as the Fundamental Principle leads to a surprising semantic conclusion such as emotivism, *this is evidence that it is false*. For a theory of meaning that aspires to fit the facts should be soporifically boring when it comes to its linguistic predictions. (This is not to say that it should contain no surprises but that the surprises should be confined to the explanatory structures and should not spill over into the empirical outputs.) A theory that “predicts” that moral judgments mean something that nobody had ever thought of until Russell arrived on the scene is a theory that is contradicted by the evidence, namely the evidence of our linguistic intuitions. This gives us a *principled* reason for preferring the Moorean *modus tollens* to the Russellian *modus ponens*. It is not just that the premisses of the Moorean *modus tollens* are more plausible than the premisses of the Russellian *modus ponens* (though they are, of course, more plausible). The Moorean premiss (R.1') asserts the kind of datum—that “good” can play a part in truth-apt sentences—that it is the business of a theory of meaning to explain. If a thesis about meaning denies such a datum—which is what Russell’s Fundamental Principle threatens to do—then that is evidence that it is false. A theoretical claim is wrong if it fails to fit the facts. A fact is not wrong if it fails to fit a theory.

15. *Morals: Methodological and Metaethical*

There are two lessons to be learned from all this. First, the general point. If a theory of meaning leads to surprising or revisionist conclusions, *that is evidence that it is false*. Thus in so far as verificationism and the Fundamental Principle suggest something as surprising as emotivism, this is evidence that they are false, not evidence that emotivism is true. But this is less conservative than it sounds. For though a theory that entails that we don’t mean what we think we mean is probably false, a theory that entails that what we mean is false or even absurd may well be true. Widespread linguistic error verges on the inconceivable, but widespread factual error is perfectly possible. It is silly to suppose that God-talk is either meaningless or non-cognitive since this flies in the face of the linguistic evidence. It is not silly (or not silly in the same way) to suppose that there is no God. In the anti-

realist struggle, error theories tend to win out against variants of non-cognitivism.³⁶

The second lesson is more Russell-specific. As we have seen, what sabotaged Russell's version of the error theory, and may have led him to abandon it, was his commitment to the Fundamental Principle. And the only Russellian argument *for* emotivism (as opposed to *against* Moorean objectivism) that we have managed to come up with likewise relied on the Fundamental Principle. But the evidence suggests that the Fundamental Principle is false. Thus a philosopher like Russell who thought that the moral facts, if any, would have to be non-natural, but could not believe in the Moorean good, would have done a lot better to adopt a *different* kind of error theory, an error theory unburdened by the Fundamental Principle and one which consequently made no difficulty in denying naturally indefinable properties. Such was the theory of J. L. Mackie.³⁷

WORKS CITED

- AYER, A. J. *Language, Truth, and Logic*. 2nd ed. London: Gollancz, 1946 (1st ed., 1935).
- . *Part of My Life*. Oxford: Oxford U. P., 1977.
- . "My Mental Development". In Lewis Hahn, ed. *The Philosophy of A. J. Ayer*. La Salle: Open Court, 1992.
- BENNETT, JONATHAN. *Locke, Berkeley, Hume: Central Themes*. Oxford, Clarendon P., 1971.
- COFFA, J. ALBERTO. *The Semantic Tradition from Kant to Carnap: to the Vienna Station*. Cambridge: Cambridge U. P., 1991.
- DENNETT, DANIEL, AND ASBJØRN STEGLICH-Petersen. *The Philosophical Lexicon*. 8th ed. Oxford: Blackwell, 1987. 2008 ed.: <https://www.philosophicallexicon.com>
- FOOT, PHILIPPA. "Utilitarianism and the Virtues". *Mind* 94, no. 374 (April 1985): 196–209.
- GALILEO GALILEI. *Dialogue Concerning the Two Chief World Systems*. 2nd ed. Trans. Stillman Drake. Berkeley: U. of California P., 1967.
- GRIFFIN, NICHOLAS, ed. *The Cambridge Companion to Bertrand Russell*. Cambridge: Cambridge U. P., 2003.
- HENLE, PAUL. "Meaning and Verifiability" (1963). In SCHILPP, ed., *The Philosophy of Rudolf Carnap*.
- LAKATOS, IMRE. *The Methodology of Scientific Research Programme; Philosophical Papers*, Vol. I. Cambridge, Cambridge U. P., 1978.
- LANGFORD, C. H. "The Notion of Analysis in Moore's Philosophy" (1942). In SCHILPP, ed., *The Philosophy of G. E. Moore*.
- LYCAN, W. G. "Moore Against the New Skeptics". *Philosophical Studies* 103 (2001): 35–53.

³⁶ See PIGDEN, "Analytic Philosophy" (2013).

³⁷ See MACKIE, "The Refutation of Morals" (1946) and *Ethics: Inventing Right and Wrong* (1977).

-
- , “Moore’s Anti-Skeptical Strategies” (2007). In NUCCETELLI AND SEAY, eds. MCGUINNESS, B. F., ed. *Ludwig Wittgenstein and the Vienna Circle: Conversations Recorded by Friedrich Waismann*. Oxford: Blackwell, 1979.
- MACKIE, J. L. “The Refutation of Morals”. *Australasian Journal of Philosophy* 24 (1946): 77–90.
- , *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin, 1977.
- MAXWELL, GROVER. “The Later Bertrand Russell: Philosophical Revolutionary”. In George Nakhnikian, ed. *Bertrand Russell’s Philosophy*. London: Duckworth, 1974.
- MOORE, G. E. *Principia Ethica*. Revised ed. Thomas Baldwin, ed. Cambridge: Cambridge U. P., 1993 (1st ed., 1903).
- , “A Reply to My Critics” (1942). In SCHILPP, ed., *The Philosophy of G. E. Moore*.
- , *Some Main Problems of Philosophy*. London: Allen & Unwin, 1953.
- , *Selected Writings*. Thomas Baldwin, ed. London: Routledge, 1993.
- NUCCELELLI, S., AND G. SEAY, eds. *Themes from G. E. Moore: New Essays in Epistemology and Ethics*. Oxford: Oxford U. P., 2007.
- PERKINS, RAY. “Bertrand Russell’s Moral Philosophy” (2019). In WAHL, ed.
- PIGDEN, CHARLES R. “Bertrand Russell: Moral Philosopher or Unphilosophical Moralism?” (2003). In GRIFFIN, ed.
- , “Desiring to Desire: Russell, Lewis and G. E. Moore” (2007). In NUCCELELLI AND G. SEAY, eds.
- , “Coercive Theories of Meaning or Why Language Should Not Matter (So Much) to Philosophy”. *Logique et Analyse* 210 (2010): 151–84.
- , “Identifying Goodness”. *Australasian Journal of Philosophy* 90 (2012): 93–109.
- , “Analytic Philosophy”. In S. Bullivant and M. Ruse, eds. *The Oxford Handbook of Atheism*. Oxford: Oxford U. P., 2013.
- , “Russell’s Moral Philosophy” (2014). *Stanford Encyclopedia of Philosophy*.
- , “Emotivism, Error and the Metaethics of Bolshevism”. *Bertrand Russell Society Bulletin*, no. 150 (Fall 2014): 19–24.
- QUINE, W. V. O. “Two Dogmas of Empiricism” (1951). Reprinted in QUINE, *Quintessence*. R. F. Gibson, ed. Cambridge, MA: Harvard U. P., 2004.
- RUSSELL, BERTRAND. “Is There an Absolute Good?” (1922). 58 in *Papers* 9.
- , *Religion and Science*. London: Home U. Library, 1935. Reissued Oxford: Oxford U. P., 1961.
- , *IMT*₂.
- , “Reply to Criticisms” (1944). In SCHILPP, ed., *The Philosophy of Bertrand Russell*; 2a in *Papers* 11.
- , *HK*.
- , *HSEP*.
- , *PfM*.
- , *LK*.
- , “Mr. Strawson on Referring” (1957). In *MPD*; 84 in *Papers* 11.
- , *MPD*.
- , *Auto*. 2.
- , *Papers* 8, 9, 10, 11.
- , RoE. C. PIGDEN, ed.
- SCHILPP, PAUL ARTHUR, ed. *The Philosophy of G. E. Moore*. (The Library of Living Philosophers.) Evanston: Northwestern U., 1942.
- , ed. *The Philosophy of Bertrand Russell*. (The Library of Living Philosophers.) Chicago: Northwestern U., 1944.
- , ed. *The Philosophy of Rudolf Carnap*. (The Library of Living Philosophers.) La Salle: Open Court, 1963.
- SOAMES, SCOTT. *Philosophical Analysis in the Twentieth Century*, Vol. 1: *The Dawn of Analysis*. Princeton: Princeton U. P., 2003.
- WAHL, RUSSELL, ed. *The Bloomsbury Companion to Bertrand Russell*. London: Bloomsbury Academic, 2019.
- , “Sense-data, Sensibilia and Percepts” (2019). In WAHL, ed.
-