BERTRAND RUSSELL: META-ETHICAL PIONEER[1].

Charles R. Pigden

University of Otago

ABSTRACT Bertrand Russell was a meta-ethical pioneer, the original inventor of both emotivism and the error theory. But why, having abandoned emotivism for the error-theory did he switch back to emotivism in the nineteen-twenties?   Perhaps he did not relish the thought that as a moralist he was a professional hypocrite.   Also Russell's version of the error theory suffers from severe defects. He commits the naturalistic fallacy and runs foul of his own and Moore's arguments against subjectivism. These defects could be repaired but only by abandoning Russell's semantics.   Russell preferred to revert to emotivism.

INTRODUCTION

In the Preface to his book on Russell[2], Mark Sainsbury proposes to leave aside Russell's work in moral philosophy, since 'in both its main phases, it is too derivative to justify a discussion of it.' He is wrong on two counts; first in supposing that Russell's work in moral philosophy had only *two* main phases (represented by 'The Elements of Ethics' (1910)[3] and *Human Society in Ethics and Politics* (1954)[4] respectively) and secondly in supposing that in the so-called second phase it is derivative. It is true, as Sainsbury states, that Russell's position in *HSEP* is 'close to

---

[1]The various volumes of *The Collected Papers of Bertrand Russell* are cited thus CPBR 1, 8 or 13 as the case may be. Paper 38 in vol. 1 would be cited thus: CPBR 1. 38. Publication details are given in the final note.

[2]Sainsbury, Mark, (1979) *Russell*, London, Routledge, p. x.

[3]Reprinted in Russell, Bertrand, (1966) *Philosophical Essays*, 2nd edn., London, Allen and Unwin, and CPBR 9. 19. All references to the reprint in *Philosophical Essays*.

[4] Russell, Bertrand, (1954) *Human Society in Ethics and Politics*, London, Allen and Unwin.

of Hume's with a dash of emotivism' but since Russell was one of the pioneers of emotivism, his position derives from a theory that Russell himself invented. It is therefore only slightly less original than the original theory itself.   As for Sainsbury's first claim, Russell's period as a pioneer of both emotivism and the error theory surely constitutes a main phase in the development of his moral philosophy, and it is this phase that I intend to discuss. Two forms of moral anti-realism have dominated the twentieth century debate: emotivism which denies that moral judgements are either true or false and the error theory which maintains that they are 'truth-apt' but false. So far as the analytic tradition is concerned, Russell invented them both. His emotivist writings anticipate those of Ayer and Stevenson (the official inventors of emotivism) by over twenty years, and he considered and rejected a version of the error theory long before J.L. Mackie published his famous 'A Refutation of Morals' in 1946.   But Sainsbury is right about one thing. 'The Elements of Ethics' *is* 'highly indebted' to G.E, Moore's *Principia Ethica.*[5] Indeed for about ten years (from 1903 till 1913) Russell was a convert to the doctrines of *Principia Ethica* differing from him only on a few minor points of detail. This meant that when he ceased to be a moral realist (that is, when he ceased to believe that moral judgements are either true or false and that some of them are true), the moral reality he rejected was that depicted by G.E, Moore, complete with non-natural properties of good and evil to be accessed by a mysterious faculty of intuition. If there was to be such a thing as moral truth, Moore was right about what it had to be. Thus in ethics, Russell remained not only a *post*-Moorean, but a *propter*-Moorean.   His anti-realist arguments were directed against a particular, and perhaps peculiar, conception of moral reality. Nowadays an anti-realist would have to argue that Moore was right about what moral judgements mean (or about what they would mean if they were construed as truth-apt) before

---

[5] Moore, G. E. (1903) *Principia Ethica,* Cambridge, Cambridge University Press. Revised edition with a substantial introduction by Thomas Baldwin, plus a cancelled preface and additional material by Moore himself (1993). All references to the revised edition.

going on to contend that, so construed, they cannot be true.[6] Russell did not see the need for such an argument.

RUSSELL THE PROTO-EMOTIVIST OR 'NO IS FROM OUGHT'

In 1913 Russell lost his faith in the Moorean good and never regained it again. Russell himself put this down to the delicate mockeries of George Santayana in *The Winds of Doctrine*. (Though he remarks that he could never be as 'bland and comfortable' without an objective good as Santayana seemed to be.[7]) The evidence confirms his recollection. For the first essay in which there are hints of emotivism is 'On the Place of Science in a Liberal Education'[8] written during or just after his reading of Santayana. (February 1913)[9]. In 'The Place of Science in a Liberal Education' Russell states that the 'kernel of the scientific outlook' (something he wants to see more widely inculcated) is 'the refusal to see our own desires, tastes and interests as affording the key to the understanding of the world.' And in the next paragraph he censures Aristotle for allowing 'himself to decide a question of fact by an appeal to aesthetico-moral considerations.' The implication would appear to be that aesthetico-moral considerations are expressions of our 'desires, tastes and interests'. Russell returned to this theme in 'Mysticism and Logic' (written in early 1914)[10] where he states that mystic emotion 'does not reveal anything about the non-

---

[6] The same can be said of J.L. Mackie in his (1946) 'The Refutation of Morals', *Australasian Journal of Psychology and Philosophy*, vol. 24 and his (1977) *Ethics: Inventing Right and Wrong,* Harmondsworth, Penguin. Accordingly many of his critics have sought to construct a non-Moorean conception of moral reality in order to sidestep his arguments. See for example Brink, David O. (1989) *Moral Realism and the Foundations of Ethics*, Cambridge, Cambridge University Press.

[7] Russell, Bertrand, (1958) *Portraits from Memory*, London, Allen and Unwin, p. 91.

[8] CPBR 12. pp. 390-397 especially 395.

[9] Russell's reading of Santayana can be dated from a letter to Goldsworthy Lowes Dickinson of 13/2/1913 This appears in Russell, Bertrand,(1967) *The Autobiography of Bertrand Russell*, vol. 1, London, Allen and Unwin, . p. 222.. For the 'The Place of Science in a Liberal Education' see CPBR. 12. p. liv

[10] CPBR 8.2

human, or the nature of the universe in general. Good and bad, and even the higher good that mysticism finds everywhere, are the reflections of our own emotions on other things not part of the substance of the things as they are in themselves. And therefore an impartial contemplation, freed from the all preoccupation with self will not judge things good or bad'.

However the clearest statement of this general idea is in 'On Scientific Method in Philosophy'[11]. I take the argument to be this. As a mere point of logic, it is obviously fallacious to infer facts from values, Is from Ought. It may be that the world *ought* to be thus-and-so or that it would be *better* if it were, but it does not follow that that is the way it is. At least it does not follow *unless* we can add an extra premise to the effect that the world, or the relevant bits of it, are as they ought to be. Now many, perhaps most, of Russell's philosophical contemporaries thought they could help themselves to just such an extra premise. Non-cognitivism affords a reason for supposing them to be wrong. If 'all ethics, however refined, remains more or less subjective', then to use ethical notions in metaphysics is 'to legislate for the universe on the basis of the present desires of men'. And there is no reason to believe that the universe corresponds to our desires. Moreover - and here I am being a little anachronistic - if non-cognitivism is true, no statement can follow from a factual claim and an evaluation unless it follows from the factual claim alone. For on the ordinary conception of logical consequence, it cannot be the case that if the premises are true the conclusion must be, since one of the premises cannot qualify as true or false (i.e. the one which asserts that the world ought to be thus-and-so). Thus the inference. even if fitted up with a missing premise, cannot be valid.   Hence 'the notion[s] of good and evil' should be 'extruded from scientific philosophy'.

Of course if aesthetico-moral judgements are not just interpretations but *mis*interpretations of the phenomena - that is if they are not truth-valueless but *false* -

---

[11] CPBR 8.4

the arguments of aesthetico-moral metaphysicians may be valid but they will all be unsound. So Russell wins either way. *Whichever* form of moral anti-realism we adopt, inferences from Ought to Is will be debarred.

Interestingly Russell had a moral motive for his crusade against moralizing metaphysics. He thought there was value in the attempt to see the world as it is without imposing our values upon it. One of the things he had against D.H. Lawrence was that he 'mistook his wishes for facts.'[12] Moreover, by the time he came to write 'On Scientific Method in Philosophy', he had come to take a dim view of moralizing in general and especially of the moral pronouncements of philosophers. Too often they seemed to him to excuse cruelty and to sanctify gutter-patriotism.

WHY BE AN EMOTIVIST?

Now all the essays I have mentioned precede the one that is generally supposed to mark Russell's conversion to emotivism (though the last only by a few days). This is 'The Ethics of Warfare'[13]. In this essay Russell announces that 'the fundamental facts in this as in all ethical questions are feelings; all that thought can do is to clarify and harmonize the expression of those feelings, and it is such clarifying and harmonizing of my own feelings that I wish to attempt in the present article.' He then goes on to such small matters as the morality of War in general and the Great War in particular. But however clear his feelings may have become as a result of this exercise, he neither clarifies nor defends the thesis that feelings are the basis of morality. What exactly did he mean by this and why did he believe it? Answers (of a sort) emerge in his replies to two philosophers, Ralph Brompton Perry of Harvard and T.E. Hulme who wrote under the soubriquet 'North Staffs'. Perry had criticised

---

[12] *Autobiography*, vol. 1, p. 53.

[13] CPRB 13.14.

'The Ethics of Warfare' whilst Hulme had delivered a vigorous attack on the lectures subsequently published as *The Principles of Social Reconstruction* under the arresting title 'The Rubbish we Oppose'[14]. Russell's 'War and Non-Resistance: A Rejoinder to Professor Perry'[15] is a polite reply to someone he respected. It is mostly devoted to rights and wrongs of warfare but Russell does state the case for non-cognitivism. This is elaborated in the more acerbic responses to 'North Staffs' (who Russell suspected of a war-for-war's-sake militarism)[16]. It is still not quite clear in these articles precisely what Russell's meta-ethic is. Is he a subjectivist, an emotivist or even, perhaps, an error theorist? *Probably* an emotivist. After all, he was no doubt aware of Moore's refutation of subjectivism and his own very similar refutation in 'The Elements of Ethics' (of which more below). And according to the error theory, as Russell himself was to formulate it, 'the emotions of approval and disapproval do not enter into the meaning of the proposition "M is good"'. The impression that you get from his replies to North Staffs is that approval and disapproval are intimately bound up with the meanings of the moral words.   But if Russell was an emotivist during the Great War, what sort of an emotivist was he? We simply cannot say.   The theory is not properly developed.

The arguments however are rather more clear. They resolve themselves into two while a third is hinted at. 1) The moral phenomena (whatever they may be) can be explained without positing moral properties. This suggests that they are ripe for Occam's razor. ('Occam's razor ... leads me to discard the notion of absolute good if ethics can be accounted for without it.')   2. Disagreements about good and evil - or more generally about basic value judgements - give us reason to doubt whether there is anything corresponding to our alleged perceptions. 'If our views on what

---

[14] This title was chosen by C.K. Ogden who seems to have enjoyed stirring the possum.

[15] CPBR 13.31.

[16] 'Mr Russell's Reply' and 'North Staff' Praise of War', (both of 1916) CPBR 13. 43 & 44.

ought to be done were to be truly rational, we ought to have a rational way of ascertaining what things ought to exist for their own account ... On [this] point no argument is possible. There can be nothing beyond an appeal to individual tastes.'[17]

EMOTIVISM AND THE MORAL PHENOMENA

The first argument is a little obscure until we know what Russell's theory is and what the relevant phenomena are supposed to be. Given the semantic cast to Russell's thought, I suspect the following. He thinks he can explain how words like 'good' can be meaningful without supposing that there is some *thing* - some property - that they mean. This can be done *either* on the assumption that they are there to express emotions *or* on the assumption that they are empty predicates. Such explanations would need to be supplemented by an account of how the institution of morality can have arisen and of why it persists. Russell did address himself to such questions in later works such as *Power*[18] and *HSEP*. But in the absence of such an account, it is not clear to me that the anti-realist Russell of 1916 has explained the moral phenomena. And if he hasn't explained the phenomena he does not have a better - that is, a more economical, and hence more clean shaven - explanation of the phenomena than the realist G.E. Moore. I suppose Russell might reply that *Moore* does not have an account of the institution of morality either - certainly not an explanation which connects the institution with properties we are alleged to apprehend. In which case Russell's property-free theory might qualify as a better explanation of the purely semantic phenomena that the two theories *do* explain.

However there is another phenomenon that Russell may have in mind, that of moral motivation. In his reply to North Staffs he says that moral arguments as to ends can only have practical efficacy 'by altering the desires or impulses of the

---

[17] CPBR 13, pp. 325-6.

[18] Russell, Bertrand, (1938) *Power: A New Social Analysis*, London, Allen and Unwin, especially chs. x & xv.

opponent not merely his intellectual judgements'. 'I cannot imagine an argument by which it could be shown that something is intrinsically good or intrinsically bad; for this reason, ethical valuations not embodying desires or impulses cannot have any [practical?] importance.' As it stands this is simply a *non sequitur*. It may be that someone's 'intellectual judgements' as to ends will only be practically efficacious if accompanied by a change in his 'desires or impulses'. (Indeed this is close to a tautology.) But it simply does not follow that a person's intellectual judgements cannot *give rise* to a change in his desires or impulses. Hence it does not follow that 'intellectual judgements' (that is, 'ethical valuations not embodying desires or impulses') 'cannot have any importance' even if we add in the seemingly irrelevant premise that what is intrinsically good or bad cannot be demonstrated by argument. No matter how people come by their 'ethical valuations' or moral beliefs, if those beliefs can beget desires, then there is no reason to suppose that they must 'embody desires or impulses' in order to be practically important (i.e. to give rise to actions). To complete the argument, Russell needs the Humean premise that genuine beliefs cannot give rise to desires without the aid of a pre-existing want[19]. He could then argue that moral beliefs, which seem to be able to do this, are not genuine beliefs, but somehow embody desires or impulses. But the Humean premise is dubious (Can't you acquire compassionate desires by meditating on another's sufferings?) and is certainly not cited in this connection.

A better way to read this rather confused argument is this. Moral judgements, specifically judgements as to ends, *motivate*. There are two possible explanations for this fact: 1) the realist or Moorean explanation that non-natural properties of goodness and badness somehow impinge upon us, giving rise to moral beliefs, which, in turn, give rise to desires, and eventually to action; and 2) the emotivist explanation that when it comes to ends, moral judgements somehow embody desires and impulses, and hence give rise to action. Of these, the latter is the simpler,

---

[19] A premise he enthusiastically endorses in the preface to *Human Society in Ethics and Politics*.

and hence to be preferred.

The problem here is that although the emotivist explanation may be the better of the two, it may not be the best since there are other explanations in the offing. Why not assume that beliefs about what is intrinsically good or evil are socially caused, and that (perhaps because of childhood conditioning) they tend to give rise to desires? This explanation does not require us to posit non-natural properties of goodness and badness and does not require us to explain away the 'propositional appearance' of moral judgements - the fact that they look like, and have usually been taken to be, statements and or propositions. The error theorist (for it is, of course, the error theory that I am proposing) can even concede that in some cases our desires can influence our basic evaluations. After all 'hot cognition' is known to occur in other cases. But the existence of 'hot' mechanisms for the formation of beliefs does not derogate from their status *as* beliefs. A mother may believe that her son is not drowned because she wants it to be true. But her belief is a belief for all that. What all this suggests is that it may be possible to explain the phenomena of ethics without recourse to an absolute good, but the explanation need not be an emotivist one.

THE ARGUMENT FROM RELATIVITY

The second argument sounds like Mackie's 'Argument from Relativity'[20]. But Russell's argument is, I think, a cut above Mackie's . If Russell were simply arguing that the diversity of moral opinion indicates that there is not really a fact of the matter to disagree about, he would be refuted by the second paragraph of his own essay 'The War and Non-Resistance; A Rejoinder to Professor Perry'. There he admits (pp. 186-187) that we cannot agree about what are undoubtedly matters of fact, such as what are the causes of the Great War and what can be done to bring it to a happy conclusion. Ideology so distorts our thinking that consensus is very hard to come by if not impossible. (This would be Marx's way of putting it not Russell's!)

---

[20] See Mackie (1977) *Ethics: Inventing Right and Wrong*, pp. 36-38.

But it would be ridiculous to suppose that the Great War did not have a determinate set of causes or that some policies rather than others were not better calculated to bring it to an end. The fact that it is difficult to agree on such topics does not prove that there are no facts to disagree about. Why then should we make this inference when it comes to morals? As Russell himself put it in 'The Elements of Ethics', p. 20, 'the difficulty of discovering the truth does not prove there is no truth to be discovered'. Moreover the 'Argument from Relativity' has been around a long time (since ancient times in fact) and by now there is a standard response on the part of moral realists (people who believe that moral judgements have a truth-value and that some of them are true). Our practical moral judgements - that this or that should be done, that this trait is a vice and that is a virtue - are derived from what might be called our ultimate or basic evaluations with the aid of factual premises. This is most obvious if we assume, for the moment, some kind of consequentialism. Why is smoking a vice? Because it undermines ones health and tends to bring on an early death which in themselves are bad things. Why is breast-feeding a good thing to do? Because it promotes the present and future health of the infant which is a good thing in itself and the precondition for other good things. If it turned out that smoking was healthy and breast-feeding harmful we would reverse our value-judgements. Now if the bulk of our moral judgements are derived from our basic evaluations with aid of factual premises, the diversity of moral opinion could be due to differences as to the facts. Hume, after enlarging on the differences between the modern French and the ancient Greeks (who disagreed about the morality of suicide), makes precisely this point. 'Have the gods forbid self-murder? An Athenian allows that it ought to be forborn. Has the Deity permitted it? A Frenchman allows that death is preferable to pain and infamy. ... the principles upon which men reason in morals are always the same; though the conclusions which they draw are often very different.'[21] This response may be effective against 18th

---

[21] See 'A Dialogue' in Hume, David, (1975) *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, 3rd edn,. Nidditch/Selby-Bigge eds., Oxford, Oxford University Press. p. 335.

Century moral sceptics, but Russell's argument is immune. For his point is that people disagree when it comes to their *basic* evaluations, about what kinds of things are intrinsically good or bad or (if, like Tolstoy, they are not consequentialists) about what kinds of things ought to be done or avoided. Russell, for instance, approves of the peaceful exercise of man's higher powers whilst 'North Staffs' (so Russell insinuates) despises such a namby-pamby ideal and thinks that strife and combat are good in themselves.

Why should this be a problem? Because it is a disagreement at the level of what are supposed to be *perceptions*. *Theoretical* differences give no cause for concern. The fact that Aristotle thinks the sun goes round the earth and that Copernicus thinks otherwise does not indicate that there is no fact of the matter waiting to be discovered. Once we transcend the observable, truths about the cosmos are hard to come by. It is not at all surprising that rational people come to different conclusions. But suppose there is some alleged realm of fact - the spirit-world, say - to be accessed by a special perceptual faculty. The mediums who claim access to this realm do not just disagree in theory. They disagree about what they claim to *perceive*, bringing back completely contradictory accounts of what they have encountered. Under the circumstances we might come to doubt whether the alleged perceptions were perceptions at all and whether the supposed realm of fact was not really a myth. Or we might come to wonder whether we had not misconstrued this entire conversational practice. Perhaps the perceptual 'reports' are not reports at all but - say - fancy ways of expressing ones mood or of influencing other people. This argument is not decisive of course. It is *possible* that there is a realm of values but that our perceptions of it are often garbled just as it is *possible* that the differing mediums have genuine but garbled perceptions of the spirit-world. So the argument does not prove that there are no objective values. But it does call their existence into question.

There is a hint - just a hint - in Russell of a related argument which falls between the Occamist argument and the Argument form Relativity. It goes like this. People disagree in their basic evaluations. So even if you think your own intuitions are correct owing to your acquaintance with the good, you must believe in the possibility of *false* intuitions, in which people wrongly perceive goodness to inhere in states which are, in fact, bad or indifferent. These mistaken intuitions are presumably due to natural causes, to upbringing, indoctrination, temperamental bias and so forth. But if *other people's* basic evaluations can be (and indeed must be) explained away in this manner, why can't the other people return the compliment and explain away your own alleged perceptions in the same way? The diversity of moral opinion - of *basic* opinion that is - suggests that real properties of goodness and badness are not needed to underwrite the phenomenology of value or to account for people's beliefs - something that even Moorean moral realists must admit when they come to the beliefs of their opponents. And if moral properties are not needed to account for people's beliefs, they are not needed at all, since they can only influence events through the medium of human action. Hence they are ripe for the razor. To escape from this argument the Moorean must either insist that his own intuitions are privileged or construct a non-naturalistic theory of moral error. He must either claim that his intuitions, unlike those of the heathen, can only be accounted for if we assume non-natural properties, or he must posit a theory of moral error in which non-natural properties play a causal role. Neither option seems particularly plausible

CRUELTY, PERSECUTION AND THE SUBJECTIVITY OF VALUE

Russell suggests another reason why people should believe in in the 'subjectivity of ethics', though this is not a reason for supposing it to be true. He wishes 'to see in the world less cruelty, persecution punishment and moral reprobation than exists at present' and 'to this end [he thinks] a recognition of the subjectivity of ethics might conduce.' This is obviously connected with his thesis that 'the claim to universality

which men associate with their ethical judgements embodies merely the impulse to persecution and tyranny'[22]. But it is not clear that this is a thought that an emotivist can coherently entertain. The idea presumably is that the belief that our enemies are wicked is more effective in stimulating homicidal fury than simple feelings of disapproval or dislike. Hence if we became aware that moral judgements merely expressed our feelings of dislike or disapproval, we would become more benign. But the very fact that conversion to the 'subjectivity of ethics' is supposed to make a difference, suggests that moral judgements do *not* mean what an emotivist must suppose them to mean. If moral judgements merely express my feelings of dislike and disapproval, how can subscribing to an expression of those feelings make me more homicidal than I was before? Or to put it the other way round, if moral judgements merely express my feelings of disapproval or dislike and I become aware of the fact, why should I become *less* murderous than I was before? Will the knowledge that moral judgements merely express my feelings make those feelings evaporate? Russell's hope that a recognition of the subjectivity of value might conduce to less cruelty only makes sense on the assumption that moral judgements embody something else besides the emotions they are alleged to express - and the most obvious candidate is a claim to objectivity. In other words, the recognition of the subjectivity of ethics can only be expected to make a difference if it involves the recognition of some sort of mistake. Which suggests not emotivism but the error theory.

RUSSELL'S ERROR THEORY

Whether we would be better off believing in the subjectivity of values or not, Russell's version of the Argument from Relativity is quite a good one. And his Occamist argument might become so if his theory were further developed. But they are not arguments *for* emotivism or non-cognitivism Rather, they are arguments *against* the intuitionism of G.E. Moore. And there are more ways than one of not

---

[22] CPBR 13, p. 326.

being a Moorean. Russell seems to have realized this. For he abandoned emotivism in the early twenties and briefly adopted what has come to be called the error theory. The theory is expressed with admirable conciseness in, 'Is There an Absolute Good'[23], which he apparently read to the Apostles on March 4th 1922.[24]. The paper remained unpublished in Russell's lifetime and only saw the light of print in 1987. On this theory, the function of moral judgements is not to express emotions (though there may be a conversational implicature that I approve of what I pronounce to be good). Moral judgements are designed to state facts. Which facts? Facts about goodness and badness. It is just that there are no such properties and no such facts and hence that moral judgements are all false. .

Russell's thesis in a nutshell is that 'good' is an empty predicate which we foist upon things 'towards which we have emotions of approval'. 'We mistake the similarity of our emotions in the presence of A, B, C, ... for the presence of a common predicate of A, B, C,' (where predicate means *property*). But how can the predicate 'good' (and here 'predicate' means predicate) be meaningful if there is no property which it means? By functioning as a definite description. "When we judge 'M is good' we mean: 'M has that predicate [property] which is common to A, B, C, ... but is absent in X, Y, Z, ...[presumably those things of which we disapprove or to which we are indifferent] It will be seen that the emotions of approval and disapproval do not enter into the meaning of the proposition 'M is good' but only into its genesis."   Since there is no such predicate or property, 'all ethical propositions are false'.

RUSSELL'S ARGUMENTS FOR THE ERROR THEORY

We have seen reason enough for Russell to adopt the error theory in his arguments

---

[23] See CPBR 9. 58. The paper is two pages long.

[24] See CPBR 9, p. 344.

of 1916. In 'Is There An Absolute Good?' he adds a few more for good measure. To begin with the theory can account for the 'propositional surface' of moral judgements. 'There seems to be no doubt that our ethical judgements claim objectivity; but this claim, to my mind, makes them all false.' Secondly, the theory can accommodate the semantic phenomenon that 'good' is meaningful (or more accurately that 'good' occurs in meaningful contexts), without supposing that there is some *thing*, some property which it means. '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 a meaning ... And it is upon this fallacy, I think, that the most apparently cogent of Moore's arguments rests.' Since an absolute good (better a Platonic property of goodness) is something any sane ontologist would want to do without, Russell's theory is to be preferred.

Besides these arguments which are developed on the first page of the paper, Russell lists five more reasons for believing his theory on the second and final page. I don't think they add much to the arguments of 1916 but two are deserving of comment.

> (1) '[The error theory] is not considered by Moore, and the arguments which he brings against the rival theories he does consider do not apply to it.' Russell does not say what these rival theories are, but I take it that he has in mind the various forms of naturalism supposedly subverted by the naturalistic fallacy and the variants of subjectivism criticised in the *Ethics* of 1912. If so, he is mistaken. Moore's arguments *do* apply to Russell's version of the error theory. As we shall see, Russell runs foul of the naturalistic fallacy and his theory suffers from the defect that has condemned subjectivism to philosophical perdition.

(2) 'It seems to be an empirical fact that the things people judge good

are the same as those towards which they have a feeling of approval'.

Is this really an empirical fact or just a fact one might arrive at by empirical means? Although the correlation between judging to be good and having the feeling of approval fails on occasion (sometimes the things we think good leave us cold whilst the things we think bad seem quite attractive) one is inclined to suspect a conceptual connection between the two. There are three plausible explanations of this correlation: a) Russell's explanation: Our moral beliefs are *caused* by our moral sentiments and are unlikely to be caused in any other way. b) The emotivist explanation (intermittently accepted by Russell himself) : Our moral beliefs *express or encapsulate* our moral feelings. Thus in the absence of such feelings there can be no moral beliefs. c) The cognitivist explanation: Our moral sentiments are *partly constituted* by our moral beliefs. What distinguishes approval from a warm feeling of liking is not some difference in phenomenological flavour but the thought that its object is good or right. Similarly what distinguishes disapproval from hatred or dislike is the thought that what we disapprove of is bad or wrong. Our moral sentiments are feelings *that*, and what follows the that-clause is the very belief that the sentiment is invoked to explain. This theory is the only theory which *both* establishes a conceptual connection between our moral beliefs and our moral sentiments *and* allows for the fact that our moral opinions sometimes leave us cold. We can have the beliefs without the sentiments but not the sentiments without the beliefs.

As will become clear, it is explanation a) that Russell needs if his arguments are to work. Explanation b) is plainly inconsistent with his theory whilst explanation c) leaves open the possibility that our moral beliefs are (sometimes) caused by contact with the good. Moreover it reduces his account of the genesis of the moral concepts

to absurdity. His claim is that the concept of 'good' arises as a result of our tendency to approve of some things and to disapprove of others. But if the emotion of approval is the feeling that X is good, we cannot appeal to this emotion to explain the concept of goodness. The concept of goodness cannot precede its own origins. So although the correlation between our moral feelings and our moral beliefs *seems* uncontroversial, it has to be taken in a rather peculiar sense if it is to sustain Russell's argument.

WHY DID RUSSELL GIVE UP THE ERROR THEORY?

To my mind, the interesting question is not 'Why did Russell adopt the error theory?' but 'Why did he give it up in favour of emotivism?'. After all, it is arguably the better theory of the two. For one thing, it accounts very neatly for that cognitive element in morals which emotivism cannot really make sense of. Russell was, of course, prone to invent good theories and then to give them up in the face of bad arguments (this is the downside of the open-mindedness on which he so prided himself). But in this case we do not even know what the arguments were, good bad, or indifferent. Let me declare my interest. I am an error theorist and I take a very dim view of emotivism and non-cognitivism generally. It seems to me the least appetizing option on the meta-ethical menu. To put it less metaphorically, whilst various forms of moral realism (both natural and non-natural) seem to me going concerns, emotivism is definitely false. Why then did Russell descend from something close to the truth into something that is clearly mistaken? Is the error theory incompatible with other parts of his philosophy? Is *his* version of the error theory subject to a peculiar defect from which other versions of the theory are free? Or am I wrong and is the error theory erroneous after all?

MORALISTS AND HYPOCRITES

One reason, I suggest, for Russell to give up the error theory was that it was just too

much to bear. Emotivism portrays moral debate as a rather sordid and manipulative business but at least it is not dishonest. The error theory converts the moralist into a purveyor of falsehoods . If he does not believe his moral pronouncements, he is a hypocrite and if does believe (or make believe) them he is a subscriber to a set of comforting falsehoods - just the sort of person that Russell despised. Russell was much addicted to moralizing (no amount of logic, he said, even if it were his own, would make him give it up[25]) and he could not bear the thought that he belonged in either category. If he accepted the error theory his vehement moral pronouncements really would be inconsistent with his official metaethic. He therefore preferred to reject it. If this is correct, then for fear of being a believer in comforting falsehoods, Russell adopted a comforting falsehood, since (in my humble opinion) the error theory is true and emotivism false.

THE ERROR THEORY AND THE NATURALISTIC FALLACY

But perhaps there are more philosophically respectable reasons for Russell to abandon his version of the error theory. Russell claims that the arguments that Moore brings against rival theories do not apply to his. I am not so sure. I think that Moore's arguments (if they are effective at all) *do* tell against Russell's version of the error theory. Departing from Russell's terminology (where 'predicate' means property) I define 'predicate ' as a word or phrase like 'good' or 'is good', and 'property' as the universal for which a referring predicate stands. Now Moore argues that goodness is a simple or unanalysable property because 'good' is a simple or unanalysable predicate. But if 'good' really were an unanalysable predicate it could not be meaningful and empty. According to Russell a predicate (or any other non-logical word) can only be meaningful but non-referring if it can be analysed as something like a definite description. That way it can play a part in a sentence, and thus can be meaningful, without having a meaning in the sense of some thing or property which it means. But according to Moore, 'good' cannot be analysed at all.

---

[25] See Russell, 'A Reply to My Critics' in Schilpp, ed. (1944) *The Philosophy of Bertrand Russell,* p. 720.

Hence it cannot be analysed as some sort of definite description. In which case it cannot be meaningful but non-referring.

Nor is this all. When Russell analyses the predicate 'good' (or better the contexts in which 'good' occurs) he sins doubly against the naturalistic fallacy. It is not just that he propounds an analysis of 'good' - the analysis he propounds is a *naturalistic* one. What 'M is good' is supposed to mean (in my mouth) is that M possesses the predicate [property] which is common to A, B, C, ... , where 'A, B, C, ...' is a list of the things that I approve of. It is fairly clear that what Russell has in mind is a list of naturalistic items. And this is no mere slip, something that might have been put right in a better formulation of the theory. For *any* analysis that Russell might propound would have to be naturalistic. In Russell's view, many words and phrases are 'incomplete symbols', expressions which can function meaningfully in the context of a larger sentence, but which do not need meanings or referents of their own in order to make sense. (Although Scott is the author of Waverly, the phrase 'the author of Waverly' could still be used meaningfully if there were no such person as Scott and the book had been composed by a committee.) But to the end of his days Russell believed that there had to be words of which this was *not* true - words which had to have referents if they were to have any meaning at all. 'There are words which are only significant because there is something that they mean, and if there were not this something, they would be empty noises not words. ... there must be such words if language is to have any relation to fact'[26]. Russell also believed in a principle of acquaintance, that every proposition that we can understand must (when fully analyzed) contain only constituents with which we are acquainted[27]. (The ultimate constituents of a proposition are of course the words which require a referent in order to be meaningful.) Finally Russell believed that we

---

[26] See Russell, Bertrand, (1959) *My Philosophical Development*, London, Allen and Unwin. pp. 240-241.

[27] *Ibid*, p. 169.

are acquainted only with items in our direct and private experience. As Grover Maxwell points out, this amounts to a form of concept empiricism.[28] For if you put these doctrines together you arrive at the claim that all understandable propositions are definable in terms drawn from our private experience. (It was the great achievement of Russell's later philosophy to escape from the prison of phenomenalism and to reconcile these apparently solipsistic doctrines with scientific realism. Propositions could be *about* quarks and gluons even though 'quarks' and 'gluons' were not among the constituents of any fully analysed proposition.) We do not experience an absolute good (since in Russell's opinion there is no such thing) so 'good' must be definable in terms of things we *do* experience. What things are these? Sensible qualities such as redness and blueness[29]. Thus 'good' becomes an empirical, and hence a naturalistic, predicate, or more accurately, a word which can be defined (in context) in terms of empirical or naturalistic predicates. Indeed, given Russell's semantic doctrines, if 'good' lacks a referent, then it *must* be given a naturalistic definition if it is to have any cognitive content at all.

Well, perhaps Moore was wrong about the naturalistic fallacy! Perhaps 'good' *can* be analysed after all! Indeed, although he may not have realized it, Russell had good reason to hope that the principal argument for the existence of a naturalistic fallacy was itself fallacious. For the Open Question Argument[30], on which Moore relies, rests on what might be called a publicity condition: B can only constitute an analysis of 'A' if the fact that 'A' and 'B' are synonymous is obvious to every competent speaker. This publicity condition leads straight to the 'Paradox of

---

[28] See Maxwell, Grover, 'The Later Bertrand Russell: Philosophical Revolutionary' in Nakhnikian, George ed. (1974) *Bertrand Russell's Philosophy*, London, Duckworth.

[29] See Russell, *My Philosophical Development*, p. 170.

[30] For the Open Question Argument see Moore (1993) *Principia Ethica*, §13.

Analysis'[31] which had it been veridical would have proved fatal to Russell's philosophical program.

Moore's argument is that 'good' cannot mean what we desire to desire, since 'Is what we desire to desire good?'is an open question. You can understand the words of which it is composed and still be in doubt about the answer. But if 'good' really *did* mean what we desire to desire, the answer to the question would be obvious - yes. Since it is *not* obvious, the *analysandum* does not mean the same as the putative *analysans*. Now this argument presupposes the publicity condition. For an analysis to be true it must be evident to everyone that the *analysans* and the *analysandum* come to the same thing. (Or at least, the equivalence must be evident once the analysis is propounded.) But this publicity condition generates the Paradox of Analysis. The Paradox states that the enterprise of analysis can only produce platitudes or falsehoods. For if the analyses produced are *true* then (by the publicity condition) they will be obvious.. But if they are not obvious then (by the publicity condition) they won't be true. Now philosophical analysis was a large part of Russell's stock in trade. Therefore Russell (like Moore himself) had a vested interest in denying the publicity condition and with it the Open Question Argument. But if the Open Question Argument fails, then the naturalistic fallacy is not a fallacy and naturalistic definitions of 'good' might turn out to be correct.

The problem here is that Russell (like Moore himself) had a vested interest in *retaining* the naturalistic fallacy as well as an interest in rejecting it. It is essential to Moore's position that 'good' cannot be given a *naturalistic* analysis. Without this his whole conception of goodness as a non-natural property goes by the board. But Russell, too, needs the naturalistic fallacy if his arguments are to succeed. He is (as I

---

[31] I owe this point to conversations with David Lewis though it is now becoming something of a commonplace. The issue is addressed in Lewis (1989) 'Dispositional Theories of Value II', *Proceedings of the Aristotelian Society Supplementary Vol.* 63, pp. 129-132.

stressed) very much a post-Moorean meta-ethicist. He assumes all along that the only form of moral realism he needs to bother about is that of of G.E.Moore. All his arguments, both for emotivism and the error theory, are based on the assumption that the only way for moral judgements to be full-bloodedly true is for 'good' to refer to a non-natural property. But if the naturalistic fallacy is not fallacious this assumption may be false. Perhaps 'good' is equivalent to a naturalistic predicate which actually *refers*. In which case moral judgements could be both truth-apt and unproblematically true. Of course, the fact that this is a theoretical possibility does not mean that it is actually the case. 'Good' might be indefinable even though Moore's arguments for the thesis are all duds. It could still be that non-natural properties are required to make moral judgements true. In which case, it could still be that moral judgements are all false since there are no such properties. But the theoretical possibility of naturalism creates a problem. In the absence of the naturalistic fallacy it would take a lot of argument to establish these claims. Russell would have to start all over again.

Perhaps the real problem lies in Russell's semantics. Russell assumes that for a non-referring predicate to make sense it must be defined in terms of words which do refer.[32] (Though this does not preclude a pyramid of incomplete symbols so long as they are grounded in empirical givens.) This forces him to develop a naturalistic analysis of 'good' (albeit an analysis which construes 'good' as an incomplete symbol) even though his argument for the error theory rests on the assumption that such an analysis is impossible.

But Russell himself would not have seen things this way. More likely he took his semantics for granted and reasoned as follows. 'If "M is good" is read as a proposition, "good" must be an empty predicate since there is no non-natural

---

[32] A Fregean might suggest that the problems arise because he trying to define sense in terms of reference.

property of goodness for "good" to refer to. This would make all moral judgements false, a consequence that I abhor. For "good" to function meaningfully as an empty predicate, it must be construed as an incomplete symbol - which means it must be subject to an a analysis. Indeed, it must be subject to a naturalistic analysis since predicates which don't refer can only be meaningful if they can be defined in terms of predicates which do. But Moore has proved that 'good' is indefinable. And it certainly can't be defined in naturalistic terms. The mistake therefore is to assume that "M is good" is a proposition. Better to interpret it as an expression in the optative mood - "Would that M were promoted!" or something of the sort. After all, this is what I used to believe in 1916, and the arguments against Moore's theory support this conclusion nearly as well as the other. Moreover there is an added bonus. If moral talk expresses my desires then it is not actually dishonest. But if moral judgements were all false, I would be little better than a professional liar.'

There is another difficulty with Russell's version of the error theory which may have led him to reject it. It runs foul of Moore's arguments against subjectivism.

SUBJECTIVISM AND THE ERROR THEORY

Moore has several arguments against the subjectivism, but the one that interests me was set out neatly by Russell himself in 'The Elements of Ethics': 'If in asserting that A is good, X meant merely to assert that A had a certain relation to himself, say of pleasing his taste in some way; and if Y in saying that A is not good, meant merely to deny that A had a like relation to himself, there would be no subject of debate between them.'[33]   This is elaborated in Moore's *Ethics* one of the two 'shilling

---

[33] See Russell (1966) *Philosophical Essays*, p. 20

shockers' of 1912[34].   On pp. 100-103, Moore mounts the following argument. Suppose, 'M is good,' means nothing more than, 'I (the speaker) approve of M,' and that 'M is bad,' means nothing more than, "I (the speaker) disapprove of M'.   Then if I say that M is good and you say that M is bad we do not contradict each other. (There is, as Russell puts it, 'no subject of debate' between us.)   For it can both be true that I approve of M whilst you disapprove.   Yet surely we *do* contradict each other.   In which case, 'M is good,' means something more than 'I (the speaker) approve of M,' and, 'M is bad,' means something more than, 'I (the speaker) disapprove of M'.   This is, of course, a hoary old chestnut familiar from first year ethics texts.   But it is pertinent to Russell's theory.   For it seems to me that Russell's theory suffers from the very same defect.

According to Russell 'M is good' (in my mouth) means that M possesses the property common to A, B, C, ... , where 'A, B, C, ... ' is a list of the things that I approve of.   But since you don't approve of the same things as I do, when you say 'M is good' what you will mean is that M has the property common to P, Q, R, ... where 'P, Q, R, ...' is the (possibly rather different) list of the things that *you* approve of.   Similarly, 'M is bad,' in my mouth, means that M possesses the property common to X, Y, Z ... , where 'X, Y, Z ...' is a list of the things I disapprove of. 'M is bad,' (in your mouth) means that M possesses the property common to F, G, H, ... , where 'F, G, H, ...' is the (possibly rather different) list of the things *you* disapprove of.

Now suppose I say M is good and you say that M is bad.   Then what I am saying is that M possesses the property common to A, B, C, ..., whilst what you are saying is that M possesses the property common to F, G, H, ... It is plain that these assertions are quite compatible even if, as Russell alleges, they are both false.   Hence

---

[34] Moore, G.E. (1912) *Ethics*, London, Williams and Norgate. The other shilling shocker was Russell's *The Problems of Philosophy*, which came out in the same year, under the same imprint and in the same series.

there is 'no subject of debate' between us. But, 'M is good,' said by me and, 'M is bad,' said by you *do* contradict each other. Hence Russell's analysis is false.

Again I would say that the problem lies in Russell's semantics. It is because he thinks predicates which don't refer must be analysed in terms of predicates which do, that he gets himself into trouble. Without this dogma, he could simply say that 'good' stands for a non-natural property but there is no such property. This would guarantee that (non-negative) judgements involving good were all false and would be consistent with the powerful arguments advanced in 1916. If asked how we can understand the word 'good' when it corresponds to nothing in our experience, Russell could reply that we *do* understand it, and that any theory which entails that we do not is simply false.

But this response would require Russell not to be Russell. If he saw the difficulties outlined in this section, he would have been much more likely to conclude that the problem lay elsewhere. If the assumption that 'good' is an empty predicate leads to absurdities, the obvious solution would be to deny that it is a predicate (i.e. something that can function as part of a proposition). Perhaps, 'M is good,' is not a proposition about M at all, but a vehicle for expressing desires. As such it is neither true nor false. In other words, if Russell had seen these difficulties, he would have had yet another reason for abandoning the error theory for emotivism.

EMOTION AND THE ERROR THEORY

My last objection applies not to Russell's analysis of 'good' but to his account of the genesis of the moral concepts.   According to Russell, we start off with emotions of approval and disapproval. 'If A, B, C, ... are the things towards which we have emotions of approval, we mistake the similarity of our emotions in the presence of A, B, C, ... for [the] perception of a common predicate [property] of A, B, C, ... To this

supposed predicate we give the name "good".' This is not a feasible scenario. For though we do have emotions of approval and disapproval, these are not raw feels or affect-programs but feelings *that*. And what follows the that-clause is the thought that the object of approval is good, right or otherwise admirable. Thus the moral sentiments are 'socially constructed' in the boring but important sense that unless your society presents you with the relevant concepts you cannot experience the emotions. If you had never heard of the right and the good you would not be able to approve, though preferring and liking would remain a possibility. If I am right about this, the moral concepts are not derived from the moral sentiments. Since moral beliefs are constituents of the moral emotions, we could not have had the emotions before we had the beliefs. (Though there were no doubt more primitive emotions which evolved into approval and disapproval as the institutions and concepts of morality were slowly created.) Russell's genesis (or genealogy) of morals is therefore defective.

This objection does not depend on defects in Russell's semantics and applies equally to Mackie's version of the error theory. Could it justify a shift from the error theory to emotivism? This depends on the brand of emotivism. If moral judgements are supposed to express and evoke approval and disapproval, the answer is no. For on this account of approval or disapproval, such theories are circular. 'X is good,' is designed to express and evoke approval. What is approval? The feeling that X is good. Clearly, this won't do. But there is one form of emotivism that may be immune to this challenge - the theory that moral judgements express and evoke *desires*. And this is precisely the theory that Russell adopted until he gave up emotivism for the rather confused theory of *HSEP*. in 1954[35]. However, I don't think these considerations can have moved him, since he seems to have

---

[35] See Russell, Bertrand (1935) *Religion and Science*, London, Home University Library, ch. 9, and Russell (1938) *Power: A New Social Analysis*, ch. xv.

conceived of the emotions of approval and disapproval as raw feels[36]. Moreover, it may be possible to revise the error theory. To be sure we must account for the genesis of the moral concepts. But this genesis need not presuppose fully-fledged emotions of approval and disapproval. Rather we would start with more primitive emotions which would gradually evolve into approval and disapproval through a judicious admixture of ideology.

CONCLUSION

Bertrand Russell was a meta-ethical pioneer. Twenty years before Barnes and Ayer he developed a version of emotivism (albeit a rudimentary one), backed it with arguments and drew some interesting consequences (No Is from Ought and the extrusion of aesthetico-moral considerations from science and philosophy.) But his arguments provide better support for the error theory than the emotivism he professed. Indeed, his version of the Argument from Relativity is superior to the one that Mackie developed thirty years later. He invented the error theory in 1922 though he derived no glory form the fact since for some reason he decided not to publish. It is true that similar theories had been advanced before, but Russell's version is the first in the analytic tradition and a considerable improvement on the nihilistic speculations of Nietzsche and Max Stirner. For Russell prefers clarity to bombast and hysteria. Indeed, it is one of the merits of his theory that it is relatively easy to see what is wrong with it. But why did Russell abandon the error theory and revert to to emotivism? After all, the error theory is arguably the better of the two! Perhaps Russell did not relish the thought that as a moralist he was a professional hypocrite. There were, however, more respectable reasons for making the switch. Russell's version of the error theory suffers from severe defects. He commits the naturalistic fallacy and runs foul of his own and Moore's arguments against subjectivism. These defects could be repaired but only by abandoning Russell's semantics. We don't know if Russell saw these problems, but if he did, he preferred

---

[36] See Russell (1966) *Philosophical Essays*, p. 27n.

another solution. Rather than giving up his semantics he gave up the error theory, and became an emotivist. The process may not have been a progress but it is certainly instructive. As always, there is much to be learned from Russell even when he is wrong.[37]

---

[37] I now give full details of the various volumes of *The Collected Papers of Bertrand Russell* that I have used.

Russell Bertrand (1983) *The Collected Papers of Bertrand Russell* , vol. 1, *Cambridge Essays 1888-1899*, Blackwell, Brink, Griffin, Rempel & Slater eds. London, Allen and Unwin.

Russell Bertrand (1985) *The Collected Papers of Bertrand Russell* , vol. 12, *Contemplation and Action 1902-1914*, Rempel, Brink and Moran eds. London, Allen and Unwin.

Russell Bertrand (1986) *The Collected Papers of Bertrand Russell* , vol. 12, *The Philosophy of Logical Atomism and Other Essays 1914-1919*, Slater ed. London, Allen and Unwin.

Russell Bertrand (1988) *The Collected Papers of Bertrand Russell* , vol. 9, *Essays on Language, Mind and Matter, 1919-1926*, Slater and Frohman eds., London, Allen and Unwin.

Russell Bertrand (1988) *The Collected Papers of Bertrand Russell* , vol. 13, *Prophecy and Dissent 1914-1916*, Rempel with Frohman, Lippincott and Moran eds. London, Unwin Hyman.

Russell Bertrand (1992) *The Collected Papers of Bertrand Russell* , vol. 6, *Logical and Philosophical Papers 1909-1913*, Slater and Frohman eds., London, Routledge.

Russell Bertrand (1994) *The Collected Papers of Bertrand Russell* , vol. 4, *Foundations of Logic 1903-05*, Urquhart with Lewis eds. London, Routledge.