

# WAS RUSSELL'S 1922 ERROR THEORY A MISTAKE?

RAY PERKINS, JR.  
Philosophy / Plymouth State U.  
Plymouth, NH, 03264-1600, USA  
PERKRRK@EARTHLINK.NET

Recent Russell scholarship has made clear the importance of Russell's contributions to ethical theory. But his provocative two-page 1922 paper, "Is There an Absolute Good?", anticipating by two decades what has come to be called "error theory", is still little known and not fully understood by students of Russell's ethics. In that little paper, never published in Russell's lifetime, he criticizes the "absolutist" view of G. E. Moore; and, with the help of his own 1905 theory of descriptions, he exposes what he takes to be the fallacy underlying Moore's (and his own earlier) arguments regarding value judgments and puts forward a new analysis which preserves the "absolutist" meaning at the cost of rendering all value judgments false. This article attempts to: (1) make clear just what Russell was doing in his little paper and how to understand it in the evolution of his meta-ethical thinking, (2) defend his 1922 theory against some recent criticisms, and (3) suggest the most likely reasons why he so quickly abandoned his new theory.

**T**hanks largely to Charles Pigden, we now know that Bertrand Russell made several innovative and important contributions to moral theory.<sup>1</sup> Almost everybody knows that Russell held, along with G. E. Moore, an objectivist and intuitionist ethics in the first decade of the last century. So too we know of Russell's subjectivist and emotivist

---

<sup>1</sup> Charles R. Pigden, ed., *Russell on Ethics* (London: Routledge, 1999). Pigden has also written "Bertrand Russell: Moral Philosopher or Unphilosophical Moralist?" in N. Griffin, ed., *The Cambridge Companion to Bertrand Russell* (Cambridge: Cambridge U. P., 2003), pp. 475–506; and "Desiring to Desire: Russell, Lewis, and G. E. Moore", in Susanna Nuccatelli and Gary Seay, eds., *Themes from G. E. Moore: New Essays in Epistemology and Ethics* (Oxford: Oxford U. P., 2007), pp. 244–69. See n.14 for his Stanford Encyclopedia article.

metaethics which replaced his earlier objectivist theory and which he held in various forms for most of the rest of his life. But few know of the curious metaethical view which he briefly held in 1922 and which has since become known as “error theory”—a theory according to which our ethical judgments are cognitively meaningful but false. In what follows I wish to examine Russell’s error theory to understand what the theory was, some current objections to it, and why he so quickly discarded it. We shall see that the new theory relies heavily on insights imported from Russell’s 1905 theory of denoting. I shall draw attention to some overlooked details of the 1922 theory that render it invulnerable to several recent lines of attack. In the course of our examination, we shall gain a clearer picture of the theory’s role in the evolution of Russell’s metaethics and also a clearer picture of what Russell himself likely saw as its shortcomings as an analysis of moral language.

Russell never published his little (two-page) paper on error theory, “Is There an Absolute Good?”, but he did present it before a special session of the secret Cambridge Apostles Society on 4 March 1922.<sup>2</sup> And he presented it, with the help of his by then famous theory of descriptions, partly as an attack on Moore and his own earlier objectivist metaethics heavily influenced by *Principia Ethica*. What Russell does in the paper, in effect, is to sketch a metaethical theory which preserves Moore’s “absolutist” insights about the nature of the *meaning* of our value language without having to be committed to an ontology of moral facts containing any Moorean non-natural properties, viz. good and bad.

Russell charges Moore (and, by implication, himself *circa* 1903) with a fallacy, viz. that the term “good”, as used by the Moorean absolutists, is not the *name* of a special property at all. Rather it is a *description*, a truncated incomplete symbol, purporting to denote such a property, and when properly analyzed, Russell tells us, “*all* propositions in which the

---

<sup>2</sup> *Russell on Ethics*, pp. 119–24; *Papers* 9: 345–6. Pigden has benefitted, as many Russell scholars have, from what appears to be the first publication of Russell’s 1922 paper in *Russell* 6 (1986): 144–9, with an introduction by Alan Ryan. Ryan speculates, based on evidence from the Russell Archives, that “probably” the paper was written for a 4 March 1922 meeting of the secret Cambridge Conversazione Society, better known as the Apostles. This is almost certainly correct, although the title of the paper is not recorded in the meeting’s minutes of which I have obtained a copy from King’s College Archives, Cambridge. Russell was in attendance and did read a paper as moderator, concerning the question “Is There Good in Goldie [Goldsworthy Lowes Dickinson]?”

---

word ‘good’ has a primary occurrence are false.”<sup>3</sup>

In order to help explain the contextual meaning of this truncated description (“good”), Russell gives what he takes to be the genesis of our notion of “good” as follows:

We have 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 perception of a common predicate of *A, B, C, ...*. To this supposed predicate we shall give the name “good”. [But] ... the predicate “good” is not to belong to anything of which we disapprove.

He continues with his explication of our value judgments:

*A, B, C, ...* are things of which we approve; *X, Y, Z, ...* are the things of which we disapprove. We judge: “There is a predicate possessed by *A, B, C, ...* but not by *X, Y, Z, ...*” To this supposed predicate, so described, we give the name “good”.  
(*Russell on Ethics*, p. 123; *Papers* 9: 344)

Thus, when we judge

(i) *M* is good

we are really saying something like:

(i′) The property<sup>4</sup> possessed by *A, B, C ...*, but not by *X, Y, Z* is instantiated by *M*,

which, on Russell’s 1905 theory of descriptions, becomes

---

<sup>3</sup> *Russell on Ethics*, p. 122; *Papers* 9: 345. Interestingly, Moore was present at the 4 March meeting, although there is no known record of his reaction. Typically, those present at the Apostles’ meetings would record their vote (Yes or No) on the question and often add a pithy comment in justification. These votes and comments are consistent with the thesis of Russell’s paper. Moore, along with Braithwaite (elected 1921) and Ramsey (also 1921), did vote “Yes” on the question. Russell voted “No”, adding “But no one is better”—a statement literally true if there is no such property as *good*. Lionel Penrose (elected 1920) also voted “No”, adding “Because he [Goldie] does not exist; though I believe he does”—a remark *a propos* of Russell’s Theory of Descriptions and his non-existent bald French monarch.

<sup>4</sup> Russell’s (and Moore’s) use of “predicate” may easily be misunderstood. He uses it to mean *non-linguistic* objects, i.e. concepts or properties.

---

(1'') There is a property (*good*)<sup>5</sup> possessed by *A, B, C ...*, but not by *X, Y, Z ...*, and which *M* instantiates.

Now, Russell claims that there is no such property, and so all our judgments of the form “*x* is good” are therefore false.<sup>6</sup>

*The role of Russell's theory of descriptions*

It's important to understand here what the theory of descriptions does for Russell as well as what it does not do. It does *not* show that (1') is false. Rather it allows us to give our statement (1) an “absolutist” meaning *even if it is false*. And the fact that we take “good” as an incomplete symbol allows us to do just that, since, on the theory of descriptions, incomplete symbols do not have to mean (denote or name) anything in order to occur meaningfully in our statements. So on this new error theory we do not have to countenance any objective Moorean properties in order to account for the absolutist meaning of our value statements.

Of course, Moore and the earlier Russell might not be completely happy with this account of the genesis of our notion of “good” precisely because it makes *good* knowable, if it is knowable at all, only indirectly *by description*. *Principia Ethica*'s metaethics was closely connected with the idea that good was cognizable directly, by acquaintance as Russell might have said. “Good” was the name of a special property.

But this *Principia Ethica* idea was, Russell later came to think, an illusion fostered by a naive semantics. In fact Russell says in 1922 that Moore's fallacy in 1903 had sprung from an unduly simplistic theory of denoting. As Russell explains:

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.

(*Russell on Ethics*, p. 123; *Papers* 9: 345)

We might reconstruct this fallacy as a complex argument for the existence of the property *good*:

---

<sup>5</sup> Russell doesn't explicitly say, but presumably he intends the description to be a *definite* description so his analysis should say “there is exactly one property”.

<sup>6</sup> *Russell on Ethics*, p. 122; *Papers* 9: 345.

---

## Argument I

- (2) “*M* is good” has meaning.
- (3) Therefore, “good” has a meaning.

## Argument II

- (3) “Good” has a meaning.
- (4) Therefore, there is a property which “good” means.

By Russell’s theory of descriptions, if we treat “good” as a description *à la* (1’) above, (3) doesn’t follow from (2) even though (2) is true. Recall that on Russell’s 1905 theory of denoting, descriptions, unlike names, do not stand for meaning components (either individuals or properties) in the corresponding expressed propositions, although they *may* have denotations.<sup>7</sup> Thus Argument I is no better than

## Argument I’

- (2’) “The present king of France is bald” has meaning.
- (3’) Therefore, “the present king of France” has a meaning.

But Russell is quite clear that “the present king of France”, like all definite descriptions, has no meaning; nor does it have a denotation.<sup>8</sup>

The lesson here is that the meaningfulness of “The so-and-so is *F*” — whether “the so-and-so” purports to denote an individual or a property—is no guarantee of there being any such entity.<sup>9</sup> So, Argument I is invalid. And thus Argument II is not cogent, i.e. its soundness, and hence its conclusion (4), is doubtful.

But a fallacious argument doesn’t have to have a false conclusion. So why did Russell come to think it untrue that there was such a property as the Moorean *good*? The answer is that, apart from not discovering any convincing arguments for the existence of such a property, there were

---

<sup>7</sup> See “On Denoting”, *LK*, pp. 55–6; *Papers* 4: 427–7. “Knowledge by Acquaintance and Knowledge by Description”, *Proceedings of the Aristotelian Society*, n.s. II (1910–11): 108–28; *ML*; *Papers* 6.

<sup>8</sup> For an analysis and defence of Russell’s account of the meaninglessness of definite descriptions, see my “Incomplete Symbols in *Principia Mathematica* and Russell’s ‘Definite Proof’”, *Russell* 31 (2011): 29–37; also “Why ‘On Denoting?’”, *ibid.*, 27 (2007): 33–9.

<sup>9</sup> In Russell’s analysis of value statements, “the so-and-so” purports to denote the property *good*, and “*F*” the property of being instantiated by *M*.

---

doubts raised by the problem of seemingly unresolvable disagreements over intrinsic value.<sup>10</sup> One might well think that if *good* were a real objective property, then disputes over whether *M* is good ought to be resolvable as most disputes in science are, by appeal to logic and experience. But at least equally important are the five reasons given by Russell in his 1922 paper for accepting the error theory.<sup>11</sup> His reasons (2)–(4) amount to the observation that since we judge “*M* is good” (or “*M* is bad”) if and only if we have the emotions of approval (or disapproval) towards *M*, and since we disagree in our ethical judgments to the same extent to which we differ in our emotions of approval and disapproval, we don't make any practical gain in assuming that there are such absolutist properties as the Moorean *good* and *bad* over and above these emotions. Hence his fifth reason—an appeal to Occam's Razor: if we don't need to assume Moorean properties to account for the *meaning* of our value judgments and for their *practical use*, we ought not to do so. Their existence is at least unlikely, and belief in them unwarranted.

*First objection: the problem of “no subject of debate”*

Pigden's main objection in *Russell on Ethics* can be stated as a *reductio ad absurdum* to the effect that Russell's error theory has the unacceptable consequence that obviously contradictory value judgments (such as “*M* is good” and “*M* is bad”) may not really contradict one another.<sup>12</sup> Pigden reminds us that Russell himself once pointed out, as his main reason for rejecting subjectivism, that it rendered our moral disagreements over *M* as mere statements of our own divergent feelings, with the absurd consequence that “there would be no subject of debate” between us. But, says Pigden, this is just what Russell's error theory allows.<sup>13</sup>

---

<sup>10</sup> See *Russell on Ethics*, p. 117 (*Papers* 13: 326), where Russell, in a 1916 letter to the *Cambridge Magazine*, replies to a critic (T. E. Hulme) and affirms that he (Russell) has abandoned belief in the objectivity of ethics in large part because he “cannot imagine any argument” for proving the truth of judgments of intrinsic value. Also see *ibid.*, p. 142, where Russell—by now (1935) a full-blown emotivist—reaffirms (from Chapter 9 of his *Religion and Science*) “the complete impossibility” of such arguments.

<sup>11</sup> *Russell on Ethics*, pp. 123–4; *Papers* 9: 346.

<sup>12</sup> *Russell on Ethics*, pp. 21–2.

<sup>13</sup> Moore also made this criticism of subjectivism, as Russell knew. (See Moore's *Ethics* [New York: Oxford U. P., 1965; 1st edn., 1912], p. 63.) Pigden appears not to have noticed that, as Russell implies in his first reason for accepting the new theory, he apparently believed that his error theory was *not* vulnerable to the “no subject of dispute”

Suppose I say “*M* is good” and you say “*M* is bad”. On Russell’s analysis, I’m saying:

- (5) There is a property (call it “*good*”) common to things (*A*, *B*, *C*, ...) that I approve of and which *M* instantiates.

You are saying:

- (6) There is a property (call it “*bad*”) common to things (*X*, *Y*, *Z*, ...) that you disapprove of and which *M* instantiates.

Pigden says that, although both (5) and (6) will be false on Russell’s theory, they may still be compatible, “For a thing might possess *both* the property common to *A*, *B*, *C*, ..., and the property common to *X*, *Y*, *Z*, ... (if there were such properties)” (p. 22, my emphasis). And he concludes:

But, “*M* is good” said by me and “*M* is bad” said by you plainly *do* contradict each other, whatever we respectively approve and disapprove of. Hence, the analysis is false. (*Russell on Ethics*, p. 22)

There are several points to make here. If we ignore Russell’s first reason for accepting his theory (see n.13), we might think that Pigden might be correct that this alleged feature of Russell’s error theory violates a primary desideratum for a satisfactory metaethical theory for the early Russell (and Moore). But this may not be a fair criticism of Russell in 1922. Russell could surely reply that his (and Moore’s) earlier account of value judgments was metaphysically and epistemically gratuitous, and his 1922 account avoids that mistake while preserving the objectivist meaning of our ordinary value talk; and that the account yields the added benefit of allowing us to see that our ethical disagreements, while real, are not genuine contradictions, much like the analysis of denotationless descriptions in Russell’s 1905 theory. When I say:

- (7) The present king of France is bald,

And you say:

---

objection. He says, “... the arguments which he [Moore] brings against the rival theories ... do not apply against it” (p. 123; *Papers* 9: 346).

---

- (8) The present king of France is not bald (where “the present king of France” has, what Russell called, primary occurrence),

we do not strictly contradict each other in the logical sense that our statements must have opposite truth values, although we certainly disagree with respect to the property attributed to a supposed monarch. *If there were* a king of France in 1905, he could not be both bald and not bald. In the case of (5) and (6), Russell might similarly say there is no strict contradiction; both statements are false. And certainly we do have a disagreement (in attitude) in our emotions towards *M*.

But do we really, in (5) and (6), have a situation similar to (7) and (8) regarding the supposed properties we attribute to *M*? Would their instantiation by *M* be impossible *were they to exist* in the same way that the king of France, were he to exist, could not possibly be both bald and not bald? Is Pigden not correct that they might both be instantiated by *M*? But if so, Russell's analysis contravenes a plain fact of our moral discourse, viz. that “*M* is good” and “*M* is bad” *do* contradict. Thus, Pigden concludes that Russell's error theory analysis is inadequate and should be rejected.

### *The problem of multiple properties*

I think Pigden's alleged refutation of Russell's analysis rests on an oversight concerning Russell's guidelines for defining the incomplete symbol, which I shall point out, but I first want to raise a related criticism: how can Russell be confident that the supposed property does not exist?<sup>14</sup> After all, it seems that all our objects of approval do share some properties, e.g. that of being in space and time. And surely they all share the property of being approved by us. There plausibly are multiple such properties common to our objects of approval, all capable of rendering our value judgments true. Thus it would seem, *pace* Russell, that judgments like (5) above would be true, not false.<sup>15</sup>

Russell's little paper seems to have anticipated this problem:

---

<sup>14</sup> Pigden raises this objection in his 2007 article “Russell's Moral Philosophy”, <http://plato.stanford.edu/entries/russell-moral/>, sec. 8.

<sup>15</sup> If we take the analysis to be providing a *definite* description for *good*, the existence of multiple properties will make all value judgments false, just as the existence of Russell and Whitehead makes both “The author of *Principia Mathematica* was mortal” and “The author of *Principia Mathematica* was not mortal” false.

---

It may be that *A, B, C, ...* will have several common predicates, but the irrelevant ones can be eliminated by the rule that the predicate “good” is not to belong to anything of which we *disapprove*.

(*Russell on Ethics*, p. 123; *Papers* 9: 345)

Let’s call this:

*Russell’s Elimination Rule 1*: The predicate *good* is not to belong to anything of which we disapprove.

It’s clear that this rule (which Pigden never mentions) will eliminate *some* candidates for “good”, e.g. the property of being in space-time, since some of our objects of disapproval would surely have this property. But unfortunately it won’t (at least not obviously) eliminate the property of being an object of our approval. All our objects of approval, and apparently none of our objects of disapproval, instantiate that property.

What is obviously needed is another elimination rule that will unambiguously describe the sort of property that Moorean absolutists took *good* to be. A close reading of Russell suggests just such a rule.

On the first page of his 1922 paper, he says that he wants an analysis of value judgments that will preserve Moore’s insight that “our ethical judgments claim objectivity.” Russell says our analysis needn’t be concerned with the relatively minor question of *good*’s alleged simplicity, but it must preserve this claimed objectivity of our value judgments. On such an analysis, “*all* propositions in which the word ‘good’ has a primary occurrence are false”—provided, he says, “*we define it [‘good’] as nearly as possible in accordance with the usage of absolutists*”.<sup>16</sup>

Here we have, I think, another, and crucially important, rule—apparently unnoticed by Pigden—for a proper analysis of *good*:

*Russell’s Elimination Rule 2*: The predicate *good* is to be an incomplete symbol defined in context *as nearly as possible in accordance with the usage of the [Moorean] absolutists*.

Put simply, it means that the symbol for *good* is to be a description which purports to denote something approximating the Moorean *good*, i.e. a property which, at least, is objective, and non-natural.

---

<sup>16</sup> *Russell on Ethics*, p. 122; *Papers* 9: 345. My italics, except for “*all*”.

---

Now let's return to Pigden's criticism and ask if the situation that he describes could actually arise in (5) or (6). Could *M* instantiate both *good* and *bad*, i.e. the special Moorean properties purportedly denoted by descriptions fashioned in accordance with Russell's two elimination rules? Recall that Pigden's criticism claims the compatibility of "good" and "bad" *in the case where the described properties are assumed to actually exist*. I think it's clear from the logic of Moore's use of "good" and "bad" in *Principia Ethica* and elsewhere that, as regards intrinsic value, it's necessarily true that *M* is good only if *M* is not bad.<sup>17</sup> So, if the special absolutist property that my "good" purports to denote were to be instantiated by *M*, then, contrary to what Pigden claims, the special absolutist property that your "bad" purports to denote could *not* also be so instantiated.

*Second objection: the problem of non-naturalness*

In his most recent article on Russell's ethics, Pigden gives less weight to his earlier (first) objection. He now claims that the main problem, and one which Russell himself felt weighty enough to cause him to abandon the error theory, is the alleged incongruity of non-naturalness with Russell's Fundamental Principle of Acquaintance.

Given Russell's theory of meaning, he ... cannot make sense of predicates that are not definable in terms of things with which we are acquainted. Thus 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" ... if it is construed as a descriptive predicate whose function is to denote a property (whether real or non-existent).<sup>18</sup>

Thus, says Pigden, to keep his error theory Russell would have been required to give up his "fundamental principle" of acquaintance.<sup>19</sup> It was something he was not willing to do.

---

<sup>17</sup> *Principia Ethica* (Cambridge: Cambridge U. P., 1903), p. 30; *Ethics*, pp. 40–3.

<sup>18</sup> Pigden, "Russell's Moral Philosophy", <http://plato.stanford.edu/entries/russell-moral/>, sec. 8.

<sup>19</sup> *Ibid.* In *My Philosophical Development*, Russell says he's long maintained a principle which he thinks still valid: "... 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).

---

This seems to misunderstand both Russell's error theory and his fundamental principle. Russell's error theory did not require him to define *good* in a way that conflicted with his fundamental principle. What it required was a definition (analysis) of "*X* is good" to be achieved with the help of a description which was, in accordance with Rules 1 and 2, suitable to denote an absolutist Moorean property, were it to exist. And that descriptive requirement would seem to be fulfilled by something like "*the Moorean property called 'good' which we claim to see as shared by all our objects of approval*". It's true that "Moorean" is to be understood as including "non-natural", but there seems to be no reason why this too cannot be understood by a description in terms of words for objects of acquaintance, or definable in terms of such words, in the same way that Russell allowed that statements about particulars unknowable by acquaintance are understood, e.g. Russell's physical objects in the *Problems*.<sup>20</sup> Properties that can't be given in acquaintance (assuming non-naturalness to be such) needn't be any less describable than physical objects that can't be so given.

*Conclusion: why Russell abandoned his new theory*

So why did Russell give up his error theory? I think the main reason has to do with Russell's notion of what constitutes a good analysis. The error theory was presented, not so much as the best analysis of moral discourse, but as the best analysis of what our ordinary moral discourse *actually means*. So if what's needed is to provide an analysis of our ordinary (absolutist) moral discourse, the analysis provided by Russell's error theory is a good one, perhaps even brilliant. But if we think, as Russell did, that good philosophical analysis should sometimes *revise our ordinary language* and tell us, not so much what we *do* mean by *X*, but what we *ought to* mean, the error theory may come up short, especially if what we ought to mean is something other than a pack of falsehoods. After all, if there are *not* any absolutist properties, then we ought not to say there *are*. For Russell, moral discourse was an important part of human life, and he must have felt more than a little discomfort about the idea of placing such importance on something that was—according to

---

<sup>20</sup> Moore had a lot to say about non-natural properties, and highly intelligible too, at least if you read Herbert Hochberg's account. See Hochberg's "Moore's Ontology and Nonnatural Properties", in E. D. Klemke, ed., *Studies in the Philosophy of G. E. Moore* (Evanston: Northwestern U. P., 1969), pp. 95–127.

---

the error theory—productive of nothing but falsehoods. A new, better analysis was needed.

The error theory did make clear that the practical business of our moral discourse is intimately connected with our emotional life and our attitudes of approval and disapproval, regardless of whether there are any absolutist moral properties. But a better metaethical theory would (1) guarantee meaningfulness for our moral language and at the same time (2) avoid the assertion of falsehoods by replacing the statemental *error-making function of the indicative mood* with an emotive *non-error-making emotive function in the optative mood*, even if it ruled out moral truth. Such a metaethical emotivism would allow Russell to keep the best of the error theory without consigning all moral talk to the realm of error. In 1922, although Russell's ideas about value judgments had already been evolving towards such a theory for nearly a decade, it would be until 1935 before he would settle on a mature version of such an emotivism in his well-known and important *Religion and Science*.<sup>21</sup>

---

<sup>21</sup> See Chapter 10 of *Religion and Science* (1935).

---