THE SUBSTITUTIONAL PARADOX IN RUSSELL'S 1907 LETTER TO HAWTREY

Bernard Linsky
Philosophy / U. of Alberta
Edmonton, ab, Canada t6g 2e5
bernard.linsky@ualberta.ca

This note presents a transcription of Russell's letter to Hawtrey of 22 January 1907 accompanied by some proposed emendations. In that letter Russell describes the paradox that he says "pilled" the "substitutional theory" developed just before he turned to the theory of types. A close paraphrase of the derivation of the paradox in a contemporary Lemmon-style natural deduction system shows which axioms the theory must assume to govern its characteristic notion of substituting individuals and propositions for each other in other propositions. Other discussions of this paradox in the literature are mentioned. I conclude with remarks about the significance of the paradox for Russell.

In the years 1905 to 1907 Bertrand Russell worked on what is now called the "substitutional theory" with its primitive notion of substituting one entity for another in a proposition, as a foundation for logic and source of a solution to the paradoxes. Russell abandoned that approach quite abruptly and returned to a logic based on propositional functions, eventually to appear as *Principia Mathematica*. Almost all of the material on the substitutional theory has remained unpublished; however, much will appear in print as the subject matter of Volume 5 of the *Collected Papers of Bertrand Russell*.[1] In recent years there has been some discussion of the substitutional theory, most prominently by

[1] Including, one may hope, a transcription of the letter which is the topic of this note.

russell: the Journal of Bertrand Russell Studies
The Bertrand Russell Research Centre, McMaster U.
n.s. 22 (winter 2002–03): 151–60
issn 0036-01631

Gregory Landini in his *Russell's Hidden Substitutional Theory.*[2] Landini reconstructs the theory and shows how the development of mathematics would have proceeded in that theory in comparison with the finally adopted approach of *Principia Mathematica.* Although there had been some speculation about Russell's reasons for abandoning the substitutional theory, it was only with work in the archives that Landini seems to have found the deciding reason, namely Russell's discovery of a paradox in the theory. Landini's book discusses variants of what he calls the "p_0 / a_0 paradox", the most striking version of which he brought to the attention of the public with a paper in *Russell* in 1989.[3] That version is found in a letter from Russell to the mathematician Hawtrey dated 22 January 1907, a copy of which is in the Russell Archives.[4] That letter serves as the frontispiece for Landini's book. The text of the letter is transcribed in I. Grattan-Guinness's *The Search for Mathematical Roots* in 2000.[5] Unfortunately, that transcription contains some errors, some pertaining to subscripts, but also including a missing negation sign. Landini, for his part, merely presents the first line of the derivation, containing the definition of the fatal p_0 and a_0, then describes the substitution that makes the trouble and then passes on. Indeed, as I will claim, Russell himself missed two subscripted "0"s in his presentation of his own paradox in this informal letter. Thus none of these sources accurately presents the details of the argument. Judy Pelham and Alasdair Urquhart have presented not only the precise instances used in the paradox, but also a free paraphrase of the argument in the letter.[6] So in all

[2] Landini, *Russell's Hidden Substitutional Theory* (New York: Oxford U. P., 1998).

[3] Landini, "New Evidence concerning Russell's Substitutional Theory of Classes", *Russell*, n.s. 9 (1989): 26–42.

[4] RA Rec. Acq. 394. The original is in the Hawtrey papers, Churchill College, Cambridge, file 10/81.

[5] *The Search for Mathematical Roots 1870–1940; Logics, Set Theories and the Foundations of Mathematics from Cantor through Russell to Gödel* (Princeton and Oxford: Princeton U. P., 2000), sec. 11.3, pp. 579–80.

[6] Pelham and Urquhart, "Russellian Propositions", in *Logic, Methodology and Philosophy of Science*, Vol. 9, ed. D. Prawitz, B. Skyrms and D. Westerståhl (Amsterdam and New York: Elsevier Science, 1994), pp. 307–26. Michael Potter also provides a loose paraphrase of the letter in *Reason's Nearest Kin: Philosophies of Arithmetic from Kant to Carnap* (Oxford: Oxford U. P., 2000), pp. 131–4. Potter dates the letter from 1909. See also the Critical Notice of Potter's book by William Demopoulos, *British Journal for the Philosophy of Science*, 52 (2001): 599–612 (at 609f.).

of this literature there is neither a precise statement of the contents of the letter nor a strict line-by-line paraphrase.

The purpose of this note is to present an accurate transcription, correct Russell's own two slips, and then give a paraphrase of the argument in a natural deduction system rather than the axiomatic formulation of logic which Russell used. I will then venture some remarks about this paradox, in particular a comparison with the original paradox of 1901 and the later paradox in Appendix B of *The Principles of Mathematics*.

The basic notion of Russell's "substitutional" theory is that of one object being substituted for another in a proposition. The notation is $p\frac{b}{a}!q$ where, as Russell says in the letter, that is to be read as meaning "p becomes q by substituting b for a". The quantifiers and variables in the theory are not typed. Restrictions arising from the very nature of substitution were to solve the paradoxes without the need for logical types. Nevertheless the variables "p" and "q" may be thought of as ranging over propositions and "a" and "b" for objects which are substituted in them, yielding new propositions.

The manuscript is clearly legible (although Grattan-Guinness thinks the year on the date is hard to make out). Russell sets out a proposition:

$$p_{\circ} \cdot = : (\exists p,\, a) : a_{\circ} \cdot = \cdot p\frac{b}{a}!q : \sim (p\frac{a_{\circ}}{a})$$

and then makes a substitution for $a_{\circ}$ in it, deriving a contradiction. The first textual problem comes from seeing what the substitution is. It is clear in the manuscript that Russell has written:

$$p_{\circ}\frac{p_{\circ}\frac{b}{a}!q}{a_{\circ}}$$

But that must be interpreted as a mistake by Russell. The substitution should rather have a subscripted "o" on the upper "a", and so should be:

$$p_{\circ}\frac{p_{\circ}\frac{b}{a_{\circ}}!q}{a_{\circ}}$$

Evidence comes from the rest of the formula, and has been noticed before. Another hand, perhaps Hawtrey's, has circled the subscript "o"

in the very last occurrence of "a_o" in the second line, and placed a question mark to the right. Russell has indeed made a mistake, not here, but earlier in the line, where he begins the argument by making a problematic substitution in the proposition p_o. It is clear from the term to the left of the second identity sign that he intends to substitute $p_o\frac{b}{a_o}!q$ for a_o in p_o. Thus the last subscripted "o" is correct, contrary to the suspicion of the earlier reader. Russell's error is repeated in the next line.

Grattan-Guinness's transcription has "p_o" rather than "p" at the end of the line beginning with "Hence", leaves out the negation sign at the beginning of the line immediately following "(1)", and is missing subscripted "o"s on "a" in two locations. Both of the "o"s are missing from "a" in a substituted formula of the form $p_o\frac{b}{a_o}!q$, the first at the end of line beginning with "Hence", the second at the end of line (3). These last two emendations are possibly silent attempts to reconcile the problem above, making the later formula consistent with the first, mistaken substitution.

These textual errors are small, but they may serve to obscure the simplicity of the actual paradox. The other source of obscurity is in the logical structure of the argument, making use as it does of Russell and Whitehead's format for derivations of conditionals that avoid repeated statements of antecedents, and some vague terminology, such as the remark that something is "shown to be involved". What follows is a statement of the steps of the argument in the general form of a natural deduction proof in the so-called "Lemmon" style, using line numbers to the left to keep track of assumptions for proofs from assumptions such as conditional proofs and *Reductio ad Absurdum*.[7] This presentation deviates from Lemmon by using Existential Instantiation (EI), (rather than EE), and in collapsing into a single step successive uses of rules such as Existential Generalization (EG), "And" Elimination (&E) and rules for identity: Identity Elimination = E, and Identity Introduction = I.

At two points in the argument a step is made that must be attributed to "axioms" of the substitutional theory, not explicitly stated in the letter, but which must none the less be present in a formal presentation

[7] E. J. Lemmon, *Beginning Logic* (Indianapolis: Hackett, 1978; 1st ed., 1965).

of the substitutional theory to which the paradox would be fatal. These are a comprehension principle in step 1, allowing for the specification of a proposition p_o by a formula that involves quantifiers over propositions and objects, and an axiom that asserts that identical propositions must have identical constituents, such as to guarantee that if $p_o\frac{b}{a_o}!q = p\frac{b}{a}!q$ then both $p = p_o$ and $a = a_o$.[8]

THE ARGUMENT IN THE LETTER TO HAWTREY

	1. $p_o = (\exists p)(\exists a)a_o = p\frac{b}{a}!q \;\&\sim(p\frac{a_o}{a})$	comprehension axiom
	2. $p_o\frac{p_o\frac{b}{a_o}!q}{a_o} = (\exists p)(\exists a)p_o\frac{b}{a_o}!q = p\frac{b}{a}!q \;\&\sim\left(p\frac{p_o\frac{b}{a_o}!q}{a}\right)$	substituting $p_o\frac{b}{a_o}!q$ for a_o in both sides of 1
3	3. $p_o\frac{p_o\frac{b}{a_o}!q}{a_o}$	assume the left-hand side of 2
3	4. $\sim\left(p\frac{p_o\frac{b}{a_o}!q}{a}\right)$	the right-hand side of 2, EI, & E
3	5. $p_o\frac{b}{a_o}!q = p\frac{b}{a}!q$	from the same instance, left conjunct
3	6. $p = p_o \;\&\; a = a_o$	5 and an axiom of the Substitutional Theory

[8] This latter axiom is seen to be plausible by realizing that the antecedent does not say only that the result of putting a_o for b in p_o is the same as the result of putting a for b in p, but rather that the proposition that q is the result of the first substitution is the same proposition as the proposition that q is the result of the second substitution. The consequent then follows by a reasonable principle of propositional identity.

3	7. $p_{o}\dfrac{p_{o}\dfrac{b}{a_{o}}!q}{a_{o}} = p\dfrac{p_{o}\dfrac{b}{a_{o}}!q}{a}$	6, = I, = E
3	8. $\sim\left(p_{o}\dfrac{p_{o}\dfrac{b}{a_{o}}!q}{a_{o}}\right)$	4, 7, = E
	9. $\sim\left(p_{o}\dfrac{p_{o}\dfrac{b}{a_{o}}!q}{a_{o}}\right)$	3–8 *Reductio*
	10. $p_{o}\dfrac{b}{a_{o}}!q = p_{o}\dfrac{b}{a_{o}}!q \;\&\sim\left(p_{o}\dfrac{p_{o}\dfrac{b}{a_{o}}!q}{a_{o}}\right)$	9, = I, & I
	11. $(\exists p)(\exists a)\; p_{o}\dfrac{b}{a_{o}}!q = p\dfrac{b}{a}!q \;\&\sim\left(p\dfrac{p_{o}\dfrac{b}{a_{o}}!q}{a}\right)$	10, EG
	12. $p_{o}\dfrac{p_{o}\dfrac{b}{a_{o}}!q}{a_{o}}$	11, 2, = E

But 12 contradicts 9! A contradiction follows from axioms of the substitutional theory.

The manuscripts contain several paradoxes for the theory, or at least different formulations of the paradox, if there is only one. One earlier attempt to reconstruct the paradox that made Russell abandon the subtitutional theory is given by Peter Hylton,[9] who constructs a contradiction for a substitutional theory that follows the paradox in Appendix B of *The Principles of Mathematics*.

That earlier paradox can be read as directed at the solution of the paradoxes in a simple theory of types.[10] It shows that if one thinks of all propositions as forming a totality, then one can also construct for

[9] Hylton, "Russell's Substitutional Theory", *Synthese*, 45 (1980): 1–31.

[10] Although André Fuhrmann argues that it should not be read that way in "Russell's Way Out of the Paradox of Propositions", *History and Philosophy of Logic*, 23 (2002): 197–213.

each class of propositions a distinct proposition saying that all propositions in that class are false, and thus getting a one-to-one correspondence between classes of propositions and propositions, leading to a contradiction. That paradox has since been of interest to several authors working on intensional logic. Landini suggests that there are important differences between the *Principles* paradox and that in the letter to Hawtrey.[11] However, in allowing for fairly unrestricted quantification over propositions with its untyped quantifiers, and assuming principles about the identity of propositions, the paradox in the letter to Hawtrey seems to involve as much of intensional logic as Hylton's, even if it is not the same argument.

How might we assess the significance for Russell of this "p_0 / a_0" paradox, or at least the version in the letter to Hawtrey? Landini quotes Russell as only saying that the paradox "pilled" the substitutional theory, and suggests that this is not to be seen as simply the same paradox returning, nor that Russell should have seen it as quite as devastating as the original.[12] Interestingly, Pelham and Urquhart read that expression as "killed", making the force of the objection quite clear. As they remark, however, this looks very much like the sort of self-referential argument that Russell produced in his first letter to Frege in 1902, and later studied as using what he called "quadratic forms". Perhaps now that it is correctly transcribed, and, closely paraphrased into a familiar idiom, it will become clear to what extent this is a new paradox, and to what extent it should be read as a very clever, and clearly quite disappointing, rediscovery of a familiar sort of paradox in what had promised to be an entirely new approach to logic. If the latter assessment is right, Russell most likely abandoned the substitutional theory as a frustrating and unsuccessful wild goose chase.

[11] *Russell's Hidden Substitutional Theory*, p. 204.

[12] Landini says, "To pill is to despoil, but not necessarily to kill" (*Russell's Hidden Substitutional Theory*, p. 234). However, the *Oxford English Dictionary* gives a slang use from 1908 meaning "to fail (a candidate) in an examination". Thus it seems that the substitutional theory is claimed to fail because of this paradox. Thanks to Nicholas Griffin who suggested that there was more to learn about the use of "pilled" in Edwardian English, and to an anonymous referee for this journal who found this entry in the *Supplement* to the first edition of the *OED*.

APPENDIX: TRANSCRIPTION OF RUSSELL'S LETTER TO HAWTREY

Bagley Wood. Jan. 22. '07

Dear Hawtrey

I forgot to send you the paradox which pilled the substitution-theory. Here it is. Put

$$p_{\circ} \,.=:\, (\exists p,\, a) : a_{\circ} \,.=.\, p\frac{b}{a}!q :\sim (p\frac{a_{\circ}}{a})$$ [where "$p\frac{b}{a}!q$" means "p becomes q by substituting b for a"]

Then

$$p_{\circ}\frac{p_{\circ}\frac{b}{a}!q}{a_{\circ}} \,.=:\, (\exists p,\, a) : p_{\circ}\frac{b}{a_{\circ}}!q \,.=.\, p\frac{b}{a}!q :\sim \left(p\frac{p_{\circ}\frac{b}{a_{\circ}}!q}{a}\right)$$

Hence

$$p_{\circ}\frac{p_{\circ}\frac{b}{a}!q}{a_{\circ}} \,.\supset :(\exists p,\, a) : p_{\circ}\frac{b}{a_{\circ}}!q \,.=.\, p\frac{b}{a}!q : p_{\circ}\frac{p_{\circ}\frac{b}{a_{\circ}}!q}{a_{\circ}} \,.\sim \left(p\frac{p_{\circ}\frac{b}{a_{\circ}}!q}{a}\right) \qquad (1)$$

$$\sim \left(p_{\circ}\frac{p_{\circ}\frac{b}{a_{\circ}}!q}{a_{\circ}}\right) \,.\supset :.\; p_{\circ}\frac{b}{a_{\circ}}!q \,.\;\; =.\, p\frac{b}{a}!\, q :\supset_{p,\,a} .\; p\frac{p_{\circ}\frac{b}{a_{\circ}}!q}{a} :.$$

$$\supset :.\, p_{\circ}\frac{p_{\circ}\frac{b}{a_{\circ}}!q}{a_{\circ}} \qquad (2)$$

$$(1) \,.\, (2) \,.\supset :.\, p_{\circ}\frac{p_{\circ}\frac{b}{a_{\circ}}!q}{a_{\circ}} :.(\exists p,\, a) : p_{\circ}\frac{b}{a_{\circ}}!q \,.=.\, p\frac{b}{a}!q :\sim \left(p\frac{p_{\circ}\frac{b}{a_{\circ}}!q}{a}\right) \qquad (3)$$

But if $p_{\circ}\frac{b}{a_{\circ}}!q$ is the same prop as $p\frac{b}{a}!q$, it seems plain we must have

$p = p_{\circ} \,.\, a = a_{\circ}$, whence $p\frac{p_{\circ}\frac{b}{a_{\circ}}!q}{a} \,.\equiv.\, p_{\circ}\frac{p_{\circ}\frac{b}{a_{\circ}}!q}{a_{\circ}}$.

Thus it is impossible that $p\frac{p_{\circ}\frac{b}{a_{\circ}}!q}{a}$ should be false while $p_{\circ}\frac{p_{\circ}\frac{b}{a_{\circ}}!q}{a_{\circ}}$ is true, which, by (3), is shown to be involved.

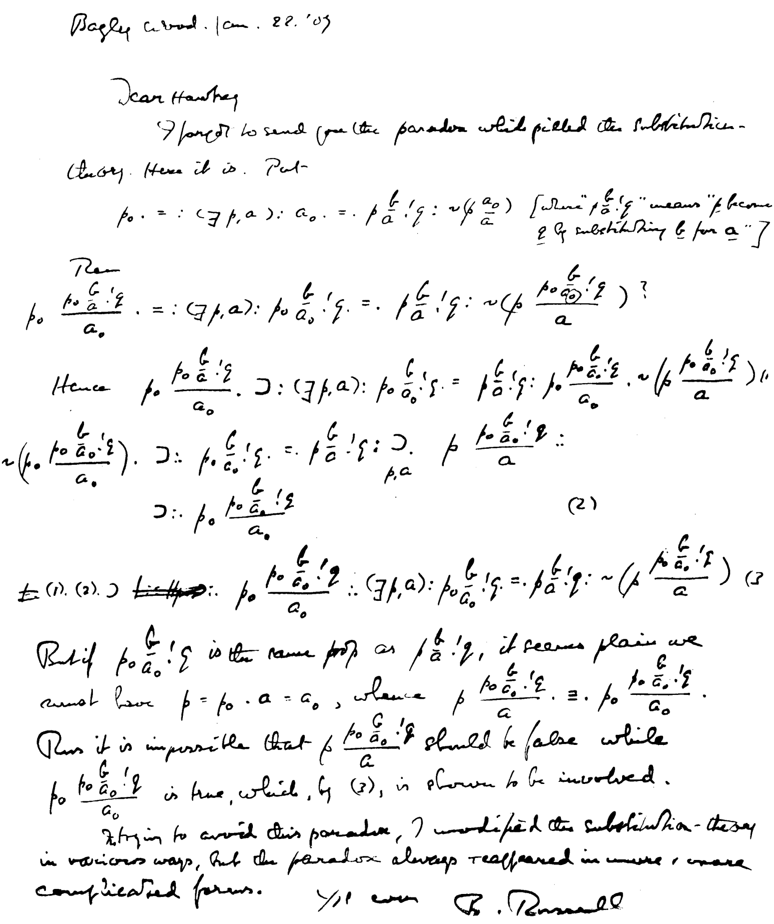
In trying to avoid this paradox, I modified the substitution-theory in various ways, but the paradox always reappeared in more and more complicated forms.

Yrs ever

B. Russell.

[Editorial note: Russell made two changes in line (3): he wrote an initial "$\vdash .$" before "(1) . (2) ." and cancelled it. After "(1) . (2) ." he wrote and cancelled "$\vdash :: \mathrm{Hp} . \supset$". This stylistic change, while not altering the logical force of the line, does make it more clearly a proof from hypotheses.]

RUSSELL'S HOLOGRAPH LETTER TO HAWTREY



Bagley Wood. Jan. 22. '07

Dear Hawtrey

I forgot to send you the paradox which pilled the substitution-theory. Here it is. Put

$p_0 . = : (\exists p, a) : a_0 . = . p\frac{b}{a}!q : \sim(p\frac{a_0}{a})$ [where "$p\frac{b}{a}!q$" means "p becomes q by substituting b for a"]

Then

$p_0 \frac{p_0\frac{b}{a}!q}{a_0} . = : (\exists p,a) : p_0\frac{b}{a_0}!q . = . p\frac{b}{a}!q : \sim(p\frac{p_0\frac{b}{a_0}!q}{a})$?

Hence $p_0\frac{p_0\frac{b}{a}!q}{a_0} . \supset : (\exists p,a) : p_0\frac{b}{a_0}!q = p\frac{b}{a}!q : p_0\frac{p_0\frac{b}{a_0}!q}{a_0} . \sim(p\frac{p_0\frac{b}{a_0}!q}{a})$ (1

$\sim(p_0\frac{p_0\frac{b}{a_0}!q}{a_0}) . \supset :: p_0\frac{b}{a_0}!q . = . p\frac{b}{a}!q : \supset_{p,a} . p\frac{p_0\frac{b}{a_0}!q}{a} ::$

$\supset :: p_0\frac{p_0\frac{b}{a_0}!q}{a_0}$ (2)

$\vdash$ (1).(2). $\supset$ [struck out] :: $p_0\frac{p_0\frac{b}{a_0}!q}{a_0} :: (\exists p,a) : p_0\frac{b}{a_0}!q = p\frac{b}{a}!q : \sim(p\frac{p_0\frac{b}{a_0}!q}{a})$ (3

But if $p_0\frac{b}{a_0}!q$ is the same prop as $p\frac{b}{a}!q$, it seems plain we must have $p = p_0 . a = a_0$, whence $p\frac{p_0\frac{b}{a_0}!q}{a} . \equiv . p_0\frac{p_0\frac{b}{a_0}!q}{a_0}$.

Thus it is impossible that $p\frac{p_0\frac{b}{a_0}!q}{a}$ should be false while $p_0\frac{p_0\frac{b}{a_0}!q}{a_0}$ is true, which, by (3), is shown to be involved.

In trying to avoid this paradox, I modified the substitution-theory in various ways, but the paradox always reappeared in more & more complicated forms.

Yrs ever B. Russell