# Hierarchical Propositions

Bruno Whittle

January 18, 2016

The notion of a proposition plays a central role in the philosophy of mind (through the notion of a propositional attitude) and the philosophy of language (through the theory of speech acts). However, as we have known for a long time, it is subject to paradoxes. A natural response is to stratify propositions into some sort of hierarchy; and, ever since Russell proposed his 'ramified theory of types' in 1908, this has been the strategy of choice here. But there is a problem that any such response must overcome, if it is to be adequate. Unfortunately, this problem does not seem to have been recognized before, let alone overcome. The upshot would seem to be that we are not in possession of an adequate hierarchical response after all. The aim of this paper, however, is to remedy this situation, by providing a hierarchical account that does overcome the problem in question.

The structure of the paper is as follows. In §1, I will introduce some paradoxes for propositions, and explain how the appeal to some sort of hierarchy would seem to be a natural and effective response to these. But in §2 I will explain the problem that this strategy faces, and how existing accounts, including Russell's, fall foul of this problem. In §§3 and 4 I will then explain how one can give a hierarchical account of propositions that overcomes the problem: §3 gives the basic idea, while §4 gives the account in full. In the appendix I give a model of the account in standard set theory.

There is something that I should make clear at the outset. This paper is about hierarchical accounts of propositions, but I do not assume that these are necessarily superior to non-hierarchical accounts. Rather, I suspect that the situation with propositions is similar to that with sets (for example). Thus, in that case, the simplest, and in many ways the most natural, response to the paradoxes is a hierarchical theory: Zermelo-Fraenkel set theory (ZF), which is hierarchical in the sense that every set belongs to a 'level', and can only contain as members things at lower levels. Despite its virtues, however, this hierarchical theory has certain limitations: for example, it would seem that according to it sets cannot play the role of extensions of predicates (at least not in general; for, since there is no set of all sets, the extension of 'is a set' cannot itself be a set, for example). The moral, it seems, is that we are going to want to have at our disposal both hierarchical and non-hierarchical set theories: since although non-hierarchical sets will be able to play roles that hierarchical sets will not be able to, it seems likely that this will come at a significant cost in terms of simplicity and naturalness. The attitude of this paper is that we should expect the situation with propositions to be similar: in this case too, hierarchical accounts (properly developed) will be simpler and more natural than non-hierarchical alternatives, but they will also face certain limitations; as a result, we are going to want to have at our disposal both sorts of account. The aim of this paper is simply to show how our hierarchical account should go.[1]

# 1 Paradoxes for Propositions and the Promise of Hierarchy

I should start by laying out some assumptions I will make in this paper. The first is that propositions are 'structured': i.e. structured in a way that more or less mirrors the structures of the sentences that express them. There are of course 'unstructured' accounts of propositions: for example, the account of them as sets of possible worlds (see, e.g., Stalnaker [1984]). But these accounts have well known—and apparently serious—problems, such as that of 'logical omniscience' (see, e.g., Stalnaker [1984: 79–99]). Thus, at least for the purposes of this paper, I will simply assume that propositions are structured. The second, more specific, assumption concerns the constituents of these propositions. Thus, there are two main broad approaches here: the 'Fregean' and the 'Russellian'. According to the former, propositions are constructed out of modes of presentation, whereas,

[1] I have mentioned sets, but one could make similar points with reference to other areas where there are paradoxes and both hierarchical and non-hierarchical responses (e.g. truth). For non-hierarchical set theories see, e.g., Forster [1995] and Maddy [1983]; for a non-hierarchical approach to propositions see, e.g., Prior [1971]. The reader should consult these works to get a sense of the extent to which going non-hierarchical requires sacrificing simplicity and naturalness.

according to the latter, they are constructed out of objects and properties.[2] In the initial discussion, I will assume a Russellian approach to propositions. However, everything I will say could easily be rephrased so as to assume a Fregean approach, if desired.[3]

I now move on to giving some paradoxes for propositions. The simplest of these is a version of Russell's paradox. Thus, suppose that $R$ is a property that applies to propositions of the form $F(a)$ (for some property $F$ and object $a$) iff $F$ does not itself apply to the proposition in question. We then consider a proposition $R(b)$ (for some object $b$) and ask: does $R$ apply to this proposition? Suppose first that it does. Then, by the definition of $R$, it must not apply to it. So $R$ does not apply to this proposition. But then—by the definition of $R$ again—$R$ *must* apply to the proposition—contradiction! And hence the paradox.[4]

For another, suppose that at time $t$ I say: any proposition that I assert at $t$ is untrue. (And suppose that I do not utter anything else at $t$.) Then I would *seem* to have asserted a proposition of the form $\forall p(A(p) \rightarrow \neg T(p))$, where $A$ is the property of being asserted by me at $t$, and $T$ is the property of being true. Call this proposition $q$. Since $q$ is the only proposition with property $A$, it would seem that it must be true iff it is not—another contradiction, and hence another paradox.

Thus, the notion of a proposition is susceptible to paradoxes. A very natural response would seem to be to propose a hierarchy, along something like the following lines. At level 0, one would start with objects that are neither properties nor propositions. At level 1, one would have properties that apply to these, and propositions that are about them. At level 2, one would have properties that apply to the things are levels 0 and 1, and propositions that are about them—and so on, with a level for each natural number (and perhaps even beyond).

Russell's 'ramified theory of types' is a more complicated response of essentially this sort. Since discussing the ramified theory of types in any detail would greatly complicate things, without changing the essential points, I will focus on the simpler response just sketched. But it will be clear to readers familiar with

[2]Note that this 'Russellian' approach is to be distinguished from Russell's ramified theory of types: the former is a general approach to propositions, while the latter is a much more specific theory, concerning the hierarchical structure of the space of propositions.

[3]For the Fregean approach to propositions, see, e.g., Frege [1892]. For the Russellian one, see, e.g., Russell [1903: 42–52] and Kaplan [1977].

[4]This paradox is similar to Russell's 'appendix B' paradox (see his [1903: 527–28]). However, I focus on the paradox in the text, rather than on that one, because the appendix B paradox involves 'classes' of propositions which seem to introduce unnecessary complications.

Russell's theory that what I will say applies to it just as it applies to the response sketched.[5]

This sort of hierarchy would seem to completely block the paradoxes. To illustrate, consider the version of Russell's paradox. On such a hierarchical account, any property $R$ will belong to a definite level of the hierarchy, and will only apply to things at lower levels. But then the closest one can get to the property $R$ that caused the problem would be a property $R^*$ such that *for some definite level n*: $R^*$ will apply to a *level n* proposition of the form $F(a)$ iff $F$ does not itself apply to it. But now any proposition of the form $R^*(b)$ (for some object $b$) will be of level greater than $n$. And so one will not be able to derive a contradiction by applying the biconditional in the definition of $R^*$ to this proposition (because that biconditional only holds for level $n$ propositions). Hence the paradox dissolves. And the other paradox that I raised is similarly blocked.[6]

[5]The main difference between the response sketched and the ramified theory of types is that the latter is 'stricter' in various ways (e.g. there can be no properties that apply both to objects and to level 1 properties). There is one point of interpretation that I should mention, however. In taking the ramified theory of types to be a response of the same sort as that sketched I am taking the former to be an account of properties and structured propositions. This clearly seems to be the correct interpretation of Russell [1908]: see, e.g., Goldfarb [1989]. But there is some debate about how exactly Whitehead and Russell [1927] is to be interpreted: see, e.g., Goldfarb [1989], Linsky [1999] and Klement [2010]. However, since I am in this paper concerned with structured propositions, I will ignore understandings of the ramified theory on which it is not an account of these (e.g. on which it is an account of linguistic entities). For more on the ramified theory of types see, e.g., Ramsey [1925], Gödel [1944], Chihara [1973], Church [1976], Anderson [1989] and Klement [2004].

[6]Thus, the paradoxes are blocked in virtue of the fact that on the sort of hierarchical account sketched properties can only apply to things at lower levels of the hierarchy. However, the sort of account sketched will also have the feature that properties and propositions can only quantify over things at lower levels. And since both of the paradoxes that I have given involve quantification (in characterizing the property $R$ I said 'for some property $F$ and object $a$'; and the version of the Liar paradox clearly involved quantification over propositions), one might wonder if one could get by *simply* with these restrictions on quantification. However, although it is possible that a hierarchical account that only had such restrictions on quantification would block the paradoxes in the text (it would depend on the details), there are alternative versions of these paradoxes that do not involve quantification. The most obvious example would be a version of Russell's paradox simply for properties: i.e. consider a property $Q$ that applies to a property $F$ iff $F$ does not apply to itself. Another example would be a version of the Liar paradox that does not involve quantification: e.g. involving a proposition $p$ of the form $\neg T(p)$, or a similar proposition constructed using a 'diagonal' function for propositions. Thus, restrictions on quantification will not be sufficient for an adequate account: one will also need the restrictions on which things a property can apply to.

However, it is not *only* that such a hierarchy will block the paradoxes: this approach would also seem to be natural in its own right, a point that can be seen as follows. To determine whether a proposition is true, one must determine whether the properties that the proposition is about apply to the objects that it is about. But on a *non*-hierarchical account one of these properties might be truth (or might depend on truth) and one of these objects might be the proposition itself. But there would then seem to be no way of determining whether the proposition is true—which is surely undesirable!

Thus, such a hierarchical account would seem to be an effective guard against paradoxes and natural in its own right. Unfortunately, such an account also faces a serious problem.

## 2 A Problem for Hierarchical Accounts

This is as follows. On a hierarchical account, no property can apply to a proposition that is itself built out of that property. But what then about propositions of the form $\neg\neg p$, for example? For here we have an operator—negation—applying to a proposition that is itself built out of that operator. But if, across the board, propositions of the form $F(F(a))$ are disallowed, then surely propositions of the form $\neg\neg p$ should be too. For why should negation be allowed to simply exempt itself from the hierarchical restrictions?

Could one argue that it should be allowed to do this because it is 'logical'? No. For our paradoxical property $R$, for example, would also seem to be logical: the notions that are used to define it, such as quantification over properties, would themselves seem to be purely logical ones.[7]

Could one instead argue that negation should be exempt from the restrictions because it is an operator, rather than a property? That is, because it has a distinctive 'pattern of combination' (it can only combine with something 'proposition shaped'). Again: no. For we could easily have raised our version of Russell's paradox not in terms of a property $R$, but in terms of an operator Ⓡ such that, for any proposition of the form $*p$ (for an operator $*$), Ⓡ$*p$ is true iff $**p$ is not. Exempting negation from the restrictions would thus seem to be ad hoc.

On the other hand, *applying* these to negation would seem to eviscerate logic. For it is surely an important and central feature of our concept of negation that it *can* be iterated. Further, the problem is by no means limited to iteration. For

[7]To spell this out: $R$ can be defined as follows. For any proposition $p$, $R(p)$ iff there is some property $F$ and object $a$ such that $p = F(a)$ and $\neg F(p)$.

example, just as one wants there to be propositions of the form $\neg\neg p$, so one wants there to be both propositions of the form $\neg(p \wedge q)$, and of the form $\neg p \wedge q$. But, on a hierarchical account, it is hard to see how this could be permissible, unless '$\neg$' and '$\wedge$' mean different things in the two cases. For, on such an account, for no properties $F$ and $G$ can there both be a proposition of the form $F(G(a))$ and one of the form $G(F(b))$. But again one surely does want an account of negation and conjunction on which there is a single negation operator $\neg$, and a single conjunction operator $\wedge$, such that there are propositions of both of these forms.[8]

We thus have a dilemma: *either* render the approach ad hoc (by simply exempting certain connectives and quantifiers from the hierarchical restrictions) *or* eviscerate logic (by not doing that!). Needless to say, neither option would seem to be very attractive.

That is the problem that hierarchical accounts face. Unfortunately, existing approaches do not seem to recognize it, let alone adequately address it. For example, Russell in [1908] (and Whitehead and Russell in [1927]) simply allow the standard connectives and quantifiers to flout the hierarchical restrictions, with no apparent recognition that there is a move being made here that requires justification. But once one *does* recognize this, it is hard not to see the resulting theory as ad hoc.[9] And commentators on Russell (and Whitehead) have seemed to be similarly unaware of the problem.[10] Further, while subsequent versions of Russell's theory, such as those of Church [1976] and Anderson [1989], represent

[8]Further, quantifiers give rise to a similar problem. For we surely want there to be propositions both of the form $\forall x \exists y H(x,y)$ and of the form $\exists y \forall x H(x,y)$, without having to say that '$\forall x$' and '$\exists y$' mean different things in the two cases. But, again, it is hard to see how this could be legitimately permissible on a hierarchical account.

[9]As I said above, there are ways in which Russell's theory differs from the simpler sort of hierarchical account sketched in §1. However, it is easy to see that none of these differences help at all with the problem that I have raised. For on Russell's theory, just as on the simpler account, there can be no propositions of the form $F(F(a))$. But then how can it be acceptable for there to be propositions of the form $\neg\neg p$ (for example)? Since no justification for this differential treatment is given, it is hard not to see the resulting theory as ad hoc.

[10]Thus, there is an extensive body of literature discussing the foundations of Russell's theory: see, e.g., the work cited in note 5. Further, large sections of this work are devoted to the question of how Russell's hierarchical restrictions might be justified (see, especially, Gödel [1944], Chihara [1973], Goldfarb [1989] and Linsky [1999]). But in none of this work does there seem to be an awareness of the fact that this question is made considerably harder by the fact that these restrictions are allowed to be flouted in certain central cases: for, once one allows exceptions in the cases of standard logical operators, it is hard to see what could justify insisting on the restrictions in every *other* case.

improvements on the original in many respects, they simply exempt standard logical operators from the hierarchical restrictions—without justification, and with no recognition that there is a problem here—in just the way that Russell did. Indeed, I know of no hierarchical account of propositions, either a version of Russell's theory, or some other approach, that recognizes or addresses this problem. It seems therefore that we are not in possession of an adequate hierarchical account after all.[11] The aim of the rest of the paper is thus to give a hierarchical account that does overcome this problem.

# 3 Solving the Problem: The Basic Idea

The basic idea behind the account that I will propose is this: rather than building propositions out of objects and properties, we instead build them out of objects and *functions*. Specifically, rather than building the proposition that John is tall, for example, out of John, together with the property of being tall, we should instead build it out of John together with the 'tallness function', i.e. the function that sends tall things to the truth value $t$, and everything else to the truth value $f$. The suggestion is that, across the board, we should construct propositions out of functions, rather than properties, in this way.[12]

How is this appeal to functions going to help with our problem? The point, essentially, is this: there are two ways in which one can 'combine' functions. Thus, given two functions $f$ and $g$, one way of combining them is simply to apply one to the other, i.e. to feed one of the functions into the other as input, so as to produce some new thing as output. But there is another way of combining

[11] Indeed, with the exception of unstructured accounts in terms of possible worlds (which, as I have noted, face serious problems), no non-hierarchical account of propositions seems to have been worked out in any detail. (Prior [1971] proposes such an account but does not so develop it.) It would seem therefore that this problem with hierarchical accounts means that we are not in possession of *any* adequate response to the paradoxes that has been worked out in any detail.

[12] I should forestall a possible confusion at this point. For Russell's ramified theory of types gives a hierarchy not only of properties and propositions but also of (what he calls) 'propositional functions'. One might thus wonder if what I am suggesting is not something that Russell has already proposed. That is not so, however. Firstly, Russell's propositional functions are not constituents of propositions (they seem rather to be the result of replacing certain constituents of propositions with variables). Secondly, if these propositional functions are literally functions at all, they are not functions to truth values (and thus nothing like the tallness function just mentioned), and they will not help with the problem raised above in anything like the way in which functions to truth values will, as will become clear once I explain that way in which functions to truth values help with the problem.

two functions: and this is to form a new function by 'composing' them. Thus, the 'composition' of $f$ and $g$, written $g \circ f$, is the function constructed out of $f$ and $g$ that, given some input $x$, works as follows: one first feeds $x$ into $f$, to produce some interim output $y$, which one then feeds into $g$, to produce the final output $z$.

If one is constructing propositions out of functions, there are then two distinct ways in which one can form a proposition out of these, because there are two distinct ways in which the functions can combine 'within' the proposition. On the one hand, there will be propositions of the form $g\langle f\rangle$, where the proposition's truth value is computed by applying $g$ to $f$.[13] But, on the other hand, there will also be propositions of the form $g \circ f\langle a\rangle$, where the truth value is computed by first applying $f$ to $a$, and then applying $g$ to the result.

This versatility of functions of course contrasts with the situation with properties: for, given two properties $F$ and $G$, one can form propositions in which one property is applied to the other (i.e. the proposition that $F$ is $G$, or the proposition that $G$ is $F$), or in which one property is applied to something built out of the other; but there is no further way in which the properties can be combined within a proposition.

But still: how does this help with the problem? The point is that, if one is giving a hierarchical account of propositions, and of functions out of which to build these, then there will be hierarchical restrictions that prohibit a function from taking as input, or as giving as output, anything at an equal or higher level to the function itself. Thus, self-application, for example, will be prohibited, and so there will be no propositions of the form $f\langle f\rangle$, whose truth value would be computed by applying $f$ to itself. On the other hand, there will be no such prohibition on self-*composition*. To see why not, suppose for example that $g$ is a function from natural numbers to natural numbers, such as the squaring function. Then—since we are giving a hierarchical account—$g$ will be at a higher level than the things that it takes as input and output (i.e. the natural numbers). But consider what we would be doing if we were to compose this function with itself—to form $g \circ g$—and then to apply this new function to some number $n$: all we would be doing would be, first, applying $g$ to $n$ to produce some new number $m$ $(= n^2)$, and then applying $g$ again to this new number to produce the final result $(= m^2 = n^4)$. That is, all we would be doing would be applying $g$ to things

[13] I use angle brackets to distinguish propositions constructed out of functions from the values of functions. That is, I will use $g\langle f\rangle$ for the proposition just mentioned, to distinguish it from the value of $g$ at $f$ (i.e. what g outputs when given $f$ as input), which I will denote $g(f)$.

lower down in the hierarchy, and producing such things as a result. Thus, there would seem to be nothing that any right-minded hierarchicalist should object to here.

So, on our hierarchical—but functional—approach, although we will not have propositions of the form $f\langle f\rangle$, we *will* have those of the form $f \circ f\langle a\rangle$ (for example). And this, it turns out, is all that we need to solve the problem.

To illustrate, consider again the issue with propositions of the form $\neg\neg p$. On the functional approach, $\neg$ will simply be the familiar truth function: i.e. the function that sends $t$ to $f$, and $f$ to $t$ (and everything else to $f$[14]). For example, let $g$ be the tallness function (which sends tall things to $t$ and everything else to $f$). The proposition that John is tall is then $g\langle\text{John}\rangle$; and the double negation of this is simply $\neg \circ \neg \circ g\langle\text{John}\rangle$, i.e. the proposition whose truth value is computed by first applying the tallness function to John, and then applying $\neg$ twice to the result, giving $t$ iff John is tall. Thus, this proposition has its truth value computed in just the way that one would expect of the double negation of John is tall. But—crucially—this computation in no way violates the hierarchical restrictions, because $\neg$ is only ever applied to things at strictly lower levels of the hierarchy (i.e. truth values), and never to anything built out of this operator. Further, the issue with negation and conjunction, as well as other similar ones, are handled in just the same way.

Thus, going functional would *seem* to solve the problem.

But—one might worry—won't this admission of self-composition allow the paradoxes to return? No. To illustrate consider again the version of Russell's paradox. This involved a property $R$ applying to propositions of the form $F(a)$ iff $F$ does not itself apply to the proposition. If one was to try to reintroduce this paradox with functions, exploiting the possibility of self-composition, one would presumably need a function $r$ such that for any proposition of the form $g\langle a\rangle$ (of some definite level of the hierarchy, say) $r \circ g\langle a\rangle$ will be true iff $g \circ g\langle a\rangle$ is not. This would seem to be the natural way of trying to reintroduce the paradox, with self-composition in place of self-application. And if one could find such an $r$, that would seem to lead to contradiction (just consider a proposition of the form $r\langle b\rangle$, and ask whether $r \circ r\langle b\rangle$ is true). Fortunately, however, it is easy to see that there will be no such $r$ within the proposed framework. For, since the truth value of $r \circ g\langle a\rangle$ is determined by *composing* $r$ with $g$—rather than *applying* $r$ to $g\langle a\rangle$—this truth value can only depend on the *truth value* of $g\langle a\rangle$ (and not on

[14]Or, strictly, everything else *of lower level than* $\neg$ to $f$. (I will sometimes omit such qualifications below.)

anything more fine-grained, as it were). But the impossibility of an $r$ as required then follows from the fact that there will be some true propositions of the form $g\langle a\rangle$ such that $g \circ g\langle a\rangle$ is true, and others such that $g \circ g\langle a\rangle$ is not. For example, let $v$ be the 'truth values' function, which sends truth values to $t$ and everything else to $f$, and let $h$ be the 'humanity' function, which sends humans to $t$ and everything else to $f$. Then consider $v\langle t\rangle$ and $h\langle\text{Obama}\rangle$. These are both true and of the requisite form. Yet $v \circ v\langle t\rangle$ is true—because the truth value of $v\langle t\rangle$ is a truth value—while $h \circ h\langle\text{Obama}\rangle$ is not—because the truth value of $h\langle\text{Obama}\rangle$ is *not* human. But then $r \circ v\langle t\rangle$ will be true iff $r \circ h\langle\text{Obama}\rangle$ is, and so it cannot be that for any proposition of the form $g\langle a\rangle$, $r \circ g\langle a\rangle$ is true iff $g \circ g\langle a\rangle$ is not. Thus, there can be no such function $r$, and self-composition does not seem to allow the paradox to re-emerge.

Of course, this argument shows merely that one way of trying to reintroduce one paradox fails. However, I will now move on to giving the proposed account in full. In the appendix I will show how to construct a model of this account in standard set theory (i.e. ZF or ZF with urelements, ZFU), from which it will follow that the account is consistent, and thus that *any* attempt to derive a contradiction from this account—i.e. any attempt to derive a paradox within it—will fail.

# 4 Solving the Problem: The Account in Full

The proposed account is thus as follows.

## 4.1 Level 0

At level 0 one starts with objects that are neither functions nor propositions. In particular, this level contains the truth values $t$ and $f$. I call the inhabitants of this level 'objects'.

## 4.2 Level 1

Level 1 is constituted as follows. To start with one has everything that one had at level 0: just because it is simplest, and less restrictive, to have a cumulative hierarchy.

But the first *new* things that one has are all the $n$-place 'unstructured' functions from objects to objects (for positive natural numbers $n$). Thus, these are functions that are not themselves built out of other functions. Rather, they are 'simple', or 'atomic', ways of going from any given $n$ objects to some other object. I will assume a 'plenitudinous' account of these: intuitively, if $X$ is a plurality of (ordered) $n+1$-tuples of objects, such that each $n$-tuple of objects is the initial segment of exactly one member of $X$, then I will assume that there is at least one unstructured function corresponding to $X$ (in the obvious way). I will not try to say more about what exactly unstructured functions *are*. Rather the proposed account will be compatible with any account of these (that is plenitudinous in the sense just described, although even this assumption could be significantly weakened if desired). For example, the proposed account will be compatible with both extensional and non-extensional accounts of these. (By an extensional account I mean one on which functions that send the same arguments to the same values are identical.)

Level 1 also contains 'structured' functions from objects to objects, constructed as follows. Thus, I assume that we have at our disposal infinitely many 'variables', which do not themselves belong to any level of the hierarchy. Further, I assume that these are divided up into infinite subclasses, one for each natural number $n$; and I refer to the variables belonging to the subclass corresponding to $n$ as 'level $n$' variables. These will of course be used to range over level $n$ of the hierarchy.[15]

Structured functions are then constructed as follows. Thus, if $f$ is an $n$-place unstructured function of level 1, and $a_1, \ldots, a_n$ are each either objects or level $\mathtt{0}$ variables, then $f\langle a_1,\ldots,a_n\rangle$ is a structured function of level 1. This can be thought of as *something like* the $n+1$-tuple of $f$ together with $a_1, \ldots, a_n$. Again, I will not try to say what exactly structured functions *are*. Rather, the proposed account will be compatible with a range of different accounts of these. All I assume is that structured functions are identical iff they have the same constituents in the same order; and that no structured function is an unstructured function.

Thus, if $a_1, \ldots, a_n$ are objects, then $f\langle a_1,\ldots,a_n\rangle$ is a constant (i.e. $\mathtt{0}$-place) function, and its value is $f(a_1,\ldots,a_n)$. On the other hand, if some $a_i$ is a variable, then $f\langle a_1,\ldots,a_n\rangle$ is a 'real' (i.e. $m$-place for positive $m$) function. For example, if $a_1$ is the only variable, then this will be a 1-place function—from objects to objects—and its value at an object $b$ will be $f(b,a_2,\ldots,a_n)$. Similarly when more than one of the $a_i$s is a variable.

[15] An alternative, perhaps 'purer', implementation of the basic idea might try to do without these variables (i.e. making do with nothing more than objects and functions). However, it simplifies things to use variables, and so, at least for the purposes of this paper, this is what I will do.

That is one way in which level 1 structured functions are constructed. But it is not the only way: for there are *also* functions constructed via 'composition' (that, after all, is why we moved to functions in the first place). Thus, let $f\langle a_1,\dots,a_n\rangle$ and $g\langle b_1,\dots,b_m\rangle$ be level 1 structured functions of the sort just constructed.[16] Then $f\langle a_1,\dots,a_{i-1},\circ g\langle b_1,\dots,b_m\rangle,a_{i+1},\dots,a_n\rangle$ will also be a structured function of level 1. $\circ$ should be thought of as an additional constituent of this function (distinct from the inhabitants of the levels of the hierarchy and the variables). Thus, this function should be thought of as *something like* the $n+2$-tuple of $f$, $a_1$, …, $a_{i-1}$, this additional constituent $\circ$, the structured function $g\langle b_1,\dots,b_m\rangle$, $a_{i+1}$, …, $a_n$.[17]

To illustrate how these functions work, let $+$ and $\times$ be unstructured addition and multiplication functions for natural numbers.[18] If $y$ is a level $0$ variable, then $+\langle\circ\times\langle 2,y\rangle,1\rangle$ will be a 1-place structured function of level 1, and its value at a number $n$ will be $+(\times(2,n),1)$ $(=2n+1)$.

More generally, if $f$ is an $n$-place unstructured function of level 1, and each of $a_1$, …, $a_n$ is either an object, a level $0$ variable, or $\circ S$, where $S$ is a structured function of level 1, then $f\langle a_1,\dots,a_n\rangle$ is a structured function of level 1. For example, $+\langle\circ\times\langle 2,\circ\times\langle 2,y\rangle\rangle,1\rangle$ is a 1-place structured function of level 1, and its value at a number $n$ is $+(\times(2,\times(2,n)),1)$ $(=4n+1)$.

That exhausts the contents of level 1. I say that an $n$-place structured function of level 1 is an $n$-place *concept* of level 1 if its values are always in $\{t,f\}$. Thus, if $g$ is the (unstructured) primeness function that sends prime numbers to $t$, and every other object to $f$, and $x$ is a level $0$ variable, then $g\langle x\rangle$ is a 1-place concept of level 1 (i.e. the concept of being prime). Further, I will say that a $0$-place concept of level 1 is a *proposition* of level 1. For example, $g\langle 23\rangle$ is a proposition of level 1 (i.e. the proposition that 23 is prime). Of course, a proposition is *true* if its value is $t$, and *false* otherwise.

[16]That is, $f$ is an $n$-place unstructured function from objects to objects, $g$ is an $m$-place such function, and $a_1$, …, $a_n$, $b_1$, …, $b_m$ are each either objects or level $0$ variables.

[17]A 'purer' implementation of the basic idea, doing without this additional constituent, is perhaps ultimately to be desired. However, for reasons of space I will not attempt such an implementation here.

[18]Thus, $+$ is a 2-place unstructured function from objects to objects such that, for any natural numbers $n$ and $m$, $+(n,m)$ is the sum of $n$ and $m$ (and any pair of objects at least one of which is not a natural number will be sent to some object that is itself not a natural number). Similarly for $\times$.

## 4.3 Connectives

It is at level 1 that we get the familiar truth functional connectives. For example, let $\neg$ be a 1-place unstructured function from objects to objects that sends $t$ to $f$, $f$ to $t$, and every other object to $f$. Then $\neg\langle\circ g\langle 23\rangle\rangle$ will be the proposition (of level 1) that 23 is not prime, while $\neg\langle\circ\neg\langle\circ g\langle 23\rangle\rangle$ will be the proposition that it is *not* not prime. Similarly, let $\wedge$ be a 2-place unstructured function from objects to objects that sends $t$, $t$ to $t$, and every other pair of objects to $f$. Then $\wedge\langle\circ g\langle 23\rangle, \circ\neg\langle\circ g\langle 24\rangle\rangle\rangle$ will be the proposition that 23 is prime and 24 is not, while $\neg\langle\circ\wedge\langle\circ g\langle 23\rangle, \circ g\langle 24\rangle\rangle\rangle$ will be the proposition that 23 and 24 are not *both* prime. And so on. Thus, as promised, we get a treatment of these connectives on which they can iterate, and combine more generally, in just the way that one would like.[19]

## 4.4 Level 2

Level 2 then stands to level 1 essentially as level 1 stood to level 0 (with just one small difference). Thus, one first has everything that one had at level 1 (because the hierarchy is cumulative); one then has all the $n$-place unstructured functions from level 1 into level 1 (for positive natural numbers $n$); and finally structured functions constructed as before. More precisely, if $f$ is an $n$-place unstructured function of level 2, but *not* an unstructured function of level 1, and each of $a_1$, $\ldots$, $a_n$ either belongs to level 1, is a level 1 variable, or is $\circ S$, where $S$ is a structured function of level 2, then $f\langle a_1, \ldots, a_n\rangle$ is a structured function of level 2. (The requirement that $f$ not belong to level 1 is the only difference from the earlier characterization of structured functions of level 1: we need this because unstructured functions of level 1 are of course not defined for all members of level 1.) As in the level 1 case, structured functions whose values are always in $\{t, f\}$ are *concepts*, and 0-place concepts are *propositions*.

[19] A reader may at this point have the following thought: why, in giving an account of functions, have I not availed myself of standard existing notation for these, specifically, Church's $\lambda$-calculus. The reason is essentially as follows. In the $\lambda$-calculus, for any variables $x$ and $y$ and formula $A$, if $\lambda x.\lambda y.A$ is wellformed then so is $\lambda y.\lambda x.A$. However, if one is giving a hierarchical account of propositions, then it is no less ad hoc to introduce operators $\lambda x$ and $\lambda y$ that can 'permute' like this than it would have been to insist from the start that standard connectives and quantifiers can iterate, permute, etc. Thus, the $\lambda$-calculus is not something one can help oneself to in giving a hierarchical account of propositions.

## 4.5 Quantifiers

It is at this point that quantification enters the picture. For example, if $x$ is level 0 variable, then level 2 will contain a corresponding existential quantifier $\exists_x$. This is a 1-place unstructured function from level 1 into level 1. To illustrate how this works, let $g$ be an $n+1$-place unstructured function from level 1 into $\{t, f\}$ (for some $n \geq 1$), and let $y_1, \ldots, y_n$ be distinct level 0 variables, each also distinct from $x$. Thus, $g\langle x, y_1, \ldots, y_n\rangle$ is an $n+1$-place concept of level 1. $\exists_x$ will then send $g\langle x, y_1, \ldots, y_n\rangle$ to an $n$-place concept of the form $h\langle y_1, \ldots, y_n\rangle$ such that, for any objects $a_1, \ldots, a_n$, $h(a_1, \ldots, a_n) = t$ iff there is *some* object $b$ such that $g(b, a_1, \ldots, a_n) = t$. For example, if $m\langle x, y\rangle$ is the is-the-mother-of concept (i.e. sending pairs of objects $c$, $d$ to $t$ iff $c$ is the mother of $d$), then $\exists_x$ will send $m\langle x, y\rangle$ to the *has*-a-mother concept $m^*\langle y\rangle$ (i.e. which sends an object $c$ to $t$ iff there is some object that is the mother of $c$). If $C$ is an $n+1$-place concept of level 1 (some $n \geq 1$) of a different form, then $\exists_x$ works similarly. On the other hand, if $C$ is a 1-place concept of level 1 of the form $g\langle x\rangle$, then $\exists_x$ will send $C$ to a truth value. For example, if $g$ is the tallness function, then $\exists_x$ will send $g\langle x\rangle$ to $t$ if some object is tall, and to $f$ otherwise. $\exists_x\langle g\langle x\rangle\rangle$ is thus the proposition that some object is tall. Similarly in the case of 1-place concepts of different forms. If $p$ is a proposition of level 1, then $\exists_x$ sends $p$ to its truth value. Finally, $\exists_x$ will send every member of level 1 that is not a concept to $f$.

Similarly for $\exists_y$ for distinct level 0 variables $y$, and for universal quantifiers $\forall_z$ for level 0 variables $z$. For example, $\exists_x\langle m\langle x, \text{John}\rangle\rangle$ is the proposition that John has a mother, while $\forall_x\langle \circ\exists_y\langle m\langle y, x\rangle\rangle$ is the proposition that everything has a mother, and $\exists_y\langle \circ\forall_x\langle m\langle y, x\rangle\rangle$ is that to the effect that there is a 'universal mother'.

As this makes clear, the proposed account allows quantifiers to combine with one another in just the way one would like. For, in the second of the three propositions just mentioned, $\exists_y$ occurs within the scope of $\forall_x$, while, in the third, it is the other way round. Of course, before moving to functions it was hard to see how such 'permutations' could be permissible on a hierarchical approach.

## 4.6 A Complication

There is, however, a complication that I should discuss at this point. This arises from the fact that we are going to have an unending hierarchy of logical operators (for example, we are going to have higher and higher level quantifiers, as we introduce higher and higher level things to quantify over). This leads to a complication because it is *not* then going to be straightforward to combine higher level operators with lower level ones: we cannot simply compose these, because the higher level operators will sometimes output things that are at too high a level for the lower level ones to take as input.

To illustrate, consider how we are going to handle propositions of the form $\exists x \neg \exists y F(x,y)$ (for distinct variables $x$ and $y$). We handled those of the form $\exists x \exists y F(x,y)$ simply by composing the level 2 functions $\exists_x$ and $\exists_y$. But we cannot now repeat the trick in the $\exists x \neg \exists y F(x,y)$ case: for when $\exists_y$ is applied to a concept of the form $f\langle x,y \rangle$ it will output, not an object, but a level 1 function; and $\neg$ cannot take this as input (because $\neg$ is *itself* such a function).

Now, if all we cared about was 'first-order' logic,[20] then this complication could be very easily solved: we could simply reconceive of $\neg$ as a level 2 function (and similarly for the other connectives);[21] and $\neg$ could then compose with our objectual quantifiers just as desired. But first-order logic is *not* all that we care about here. The aim is rather a completely general account of propositions, and so of the whole range of logical operators they can contain.

A better solution would thus seem to be as follows. When we introduce a new level of logical operators (e.g. when we introduce our 'objectual' quantifiers $\exists_x$, $\forall_y$ etc. for level $\mathtt{0}$ variables $x$, $y$ etc.) we must also introduce 'raising' functions. That is, functions that 'raise' our previously introduced (lower level) logical operators, so as to allow them to compose with the higher level operators. For example, one of these raising functions will be a 2-place level 2 function $R$ such that, when $\neg$ is 'raised' using it, the result can combine in just the way that one would like with the level 2 logical operators (i.e. $\exists_x$, $\forall_y$ etc.). More precisely, if $Z$ is a level 1 variable, then $R\langle \neg, Z \rangle$ will be a 1-place concept of level 2 that works as a level 2 version of negation,[22] and which can thus straightforwardly compose with other level 2 operators. Thus, using $R$, we can straightforwardly handle propositions of the form $\exists x \neg \exists y F(x,y)$: these will have the form

[20] That is, standard quantificational logic, the basic terms of which are: names of objects, function symbols for functions from objects to objects, predicates of objects, quantifiers and variables over objects, and truth functional connectives.

[21] Thus, $\neg$ would send an $n$-place level 1 concept of the form $g\langle x_1, \ldots, x_n \rangle$ to its 'complement', i.e. a concept $h\langle x_1, \ldots, x_n \rangle$ that sends objects $a_1, \ldots, a_n$ to $t$ iff $g\langle x_1, \ldots, x_n \rangle$ sends them to $f$ (and similarly in the cases of level 1 concepts of other forms; $\neg$ would send anything that is not a level 1 concept to $f$). Other connectives could be similarly reconceived.

[22] That is, $R\langle \neg, Z \rangle$ works just like the version of negation described in the previous note: sending level 1 concepts to their 'complements' (and sending anything that is not a level 1 concept to $f$).

$\exists_x \langle \circ R \langle \neg, \circ \exists_y \langle f \langle x, y \rangle \rangle \rangle \rangle$ (and so the truth value is computed by first applying $\exists_y$ to $f \langle x, y \rangle$, then applying $R\langle \neg, Z \rangle$ to the result, and then applying $\exists_x$ to the result of that). More generally, such raising functions will allow lower level operators to combine with higher level ones in just the way that one would like.[23]

The picture that emerges is as follows. We have a hierarchy of logical operators, and when we introduce a new level of these, we also need to 'raise' our old operators so they can combine properly with the new ones. Fortunately, it is straightforward to implement this way around the complication within the proposed account—i.e. without any sort of exceptions to the hierarchical restrictions. Indeed, it bears emphasis that the raising functions discussed above are not an *addition* to the previously described hierarchy. Rather, their existence follows from that previous description. (For example, the level 2 function $R$ discussed above belongs to level 2 in virtue of the description of §4.4; specifically, in virtue of the fact that level 2 contains all unstructured functions from level 1 into level 1.)

I should also stress how different this treatment of logical operators is from one on which such operators cannot iterate or permute. For on the proposed treatment, when one introduces a new level of logical operators, one must introduce new raising functions to allow one's previously introduced logical operators to combine with these. Once these raising functions have been introduced, however, one's logical operators (new and old) can combine just as desired. And one need only introduce more raising functions when one introduces what are—by anyone's lights—new logical operators (e.g. quantifiers over a new, 'higher', range of things). This is of course *very* different from a treatment on which one must introduce a new (higher level) logical operator every time one simply wants to iterate one of those one has already introduced.

That completes the account of levels 0–2. For each natural number $n \geq 2$, there is a level $n+1$ that stands to level $n$ just as level 2 stands to level 1. One could also, if desired, straightforwardly extend the hierarchy to transfinite levels. As before, a structured function whose values are always in $\{t, f\}$ is a *concept*, and a 0-place concept is a *proposition*.

[23]Specifically, there will be a distinct raising function $R_{n,m}$ for each pair $n$, $m$ with $n \geq 1$ and $m \geq 2$. Thus, $R_{n,m}$ will in effect turn the $n$-place logical operators of level less than $m$ into level $m$ operators, so as to allow them to compose with other level $m$ operators (such as quantifiers over level $m-1$).

This account will of course block paradoxes such as those of §1 in just the way described there. For example, because any function can only apply to things at lower levels, there can be no function $r$ such that for *any* proposition of the form $f\langle a\rangle$, $r$ sends this to $t$ iff the function $f$ does not itself send it to $t$. The closest one can get is a function $r^*$ such that *for some definite level $n$*: for any proposition *of level $n$* of the form $f\langle a\rangle$, $r^*$ will send this to $t$ iff $f$ does not itself send it to $t$. But, since propositions of the form $r^*\langle b\rangle$ will not themselves belong to level $n$, the paradox will be blocked.

As I noted in the introduction, hierarchical accounts of propositions—like hierarchical accounts of sets, for example—will have certain limitations. For example, on the account that I have given, no proposition can quantify over absolutely all functions, or absolutely all concepts or propositions. Rather, a proposition can quantify only over those functions, concepts or propositions that occur at lower levels than the proposition itself. For this sort of reason, we are presumably also going to want to develop non-hierarchical accounts of propositions. Nevertheless, the hierarchical account that I have proposed is very simple and naturally motivated—virtues it seems unlikely that a non-hierarchical account will possess to anything like the same extent. Thus, we are presumably ultimately going to want to have both sorts of account at our disposal. What I hope to have shown, then, is what our hierarchical account should look like.[24]

# Appendix

In this appendix I construct a model of the proposed account in standard set theory. In particular, in ZFU. I write $\mathbb{N}$ for the set of natural numbers $\{0,1,2,\dots\}$, and $\mathbb{N}^+$ for $\mathbb{N}-\{0\}$. For simplicity, I assume that there exist the following infinite, and pairwise disjoint, sets of urelements: $\mathrm{LEV}_0$, $\mathrm{VAR}_0$, $\mathrm{VAR}_1$, ... (i.e. a set $\mathrm{VAR}_n$ for each $n \in \mathbb{N}$). I assume also that there is some urelement $\circ$ not in any of these sets, and that the truth values $t$ and $f$ are in $\mathrm{LEV}_0$. $\mathrm{LEV}_0$ is the model of level 0 of the hierarchy.

Let $X$ and $Y$ be non-empty sets. By an $n$-place *set-theoretic function* from $X$ to $Y$ ($n \in \mathbb{N}^+$) I mean a set $Z$ of ordered $n+1$-tuples of members of $X \cup Y$ such that: the first $n$ members of each member of $Z$ are in $X$, and the last member is in $Y$; and each $n$-tuple of members of $X$ is the initial segment of exactly one member of $Z$.

[24] [Acknowledgements.]

$\mathrm{UF}_1 = \{f : f$ is an $n$-place set-theoretic function from $\mathrm{LEV}_0$ to $\mathrm{LEV}_0$ for some $n \in \mathbb{N}^+\}$. $\mathrm{UF}_1$ is the model of the class of unstructured functions of level 1.

The model of the class of structured functions of level 1 is as follows. I use $[\![a_1, \ldots, a_n]\!]$ for the $n$-tuple of $a_1, \ldots, a_n$ in that order. $\mathrm{SF}_1$ is defined recursively as follows:

> if $f$ is an $n$-place member of $\mathrm{UF}_1$ (for some $n \in \mathbb{N}^+$), and each $a_i$ is either a member of $\mathrm{LEV}_0 \cup \mathrm{VAR}_0$ or $[\![\circ, g]\!]$ for some $g \in \mathrm{SF}_1$, then $[\![f, a_1, \ldots, a_n]\!] \in \mathrm{SF}_1$.

$\mathrm{LEV}_1 = \mathrm{LEV}_0 \cup \mathrm{UF}_1 \cup \mathrm{SF}_1$. $\mathrm{LEV}_1$ is of course the model of level 1 of the hierarchy.

The model of level 2 is as follows. $\mathrm{UF}_2 = \mathrm{UF}_1 \cup \{f : f$ is an $n$-place set-theoretic function from $\mathrm{LEV}_1$ to $\mathrm{LEV}_1$ for some $n \in \mathbb{N}^+\}$. $\mathrm{SF}_2$ is then defined:

(i) $\mathrm{SF}_1 \subseteq \mathrm{SF}_2$; and

(ii) if $f$ is an $n$-place member of $\mathrm{UF}_2 - \mathrm{UF}_1$ (for some $n \in \mathbb{N}^+$), and each $a_i$ is either a member of $\mathrm{LEV}_1 \cup \mathrm{VAR}_1$ or $[\![\circ, g]\!]$ for some $g \in \mathrm{SF}_2$, then $[\![f, a_1, \ldots, a_n]\!] \in \mathrm{SF}_2$.

$\mathrm{LEV}_2 = \mathrm{LEV}_0 \cup \mathrm{UF}_2 \cup \mathrm{SF}_2$ is the model of level 2 of the hierarchy.

The models of subsequent levels are defined similarly. Thus, the model of level $m+1$ for $m \in \mathbb{N}$ with $m \geq 2$ is as follows. $\mathrm{UF}_{m+1} = \mathrm{UF}_m \cup \{f : f$ is an $n$-place set-theoretic function from $\mathrm{LEV}_m$ to $\mathrm{LEV}_m$ for some $n \in \mathbb{N}^+\}$. $\mathrm{SF}_{m+1}$ is defined as follows:

(i) $\mathrm{SF}_m \subseteq \mathrm{SF}_{m+1}$; and

(ii) if $f$ is an $n$-place member of $\mathrm{UF}_{m+1} - \mathrm{UF}_m$ (for some $n \in \mathbb{N}^+$), and each $a_i$ is either a member of $\mathrm{LEV}_m \cup \mathrm{VAR}_m$ or $[\![\circ, g]\!]$ for some $g \in \mathrm{SF}_{m+1}$, then $[\![f, a_1, \ldots, a_n]\!] \in \mathrm{SF}_{m+1}$.

$\mathrm{LEV}_{m+1} = \mathrm{LEV}_0 \cup \mathrm{UF}_{m+1} \cup \mathrm{SF}_{m+1}$.

For $n \in \mathbb{N}^+$, I write $\mathrm{VAR}_{<n}$ for $\bigcup_{m<n} \mathrm{VAR}_m$. I define the notion of a member of $\mathrm{VAR}_{<n}$ being 'free' in a member of $\mathrm{SF}_n$ as follows. Let $x \in \mathrm{VAR}_{<n}$ and $f \in \mathrm{SF}_n$. For some $g \in \mathrm{UF}_n$ and $a_1, \ldots, a_m$ ($m \in \mathbb{N}^+$), $f = [\![g, a_1, \ldots, a_m]\!]$. $x$ is *free* in $f$ if: for some $i$ ($1 \leq i \leq m$), $a_i$ is $x$; or $a_i$ is $[\![\circ, h]\!]$ and $x$ is *free* in $h$. $f \in \mathrm{SF}_n$ is *$r$-place* if exactly $r$ members of $\mathrm{VAR}_{<n}$ are free in $f$.

The 'values' of members of $\mathrm{SF}_n$ are defined as follows. I write $\mathrm{VAR}_{\mathbb{N}}$ for $\bigcup_{n\in\mathbb{N}} \mathrm{VAR}_n$ and $\mathrm{LEV}_{\mathbb{N}}$ for $\bigcup_{n\in\mathbb{N}} \mathrm{LEV}_n$. An *assignment* is a 1-place set-theoretic function $A$ from $\mathrm{VAR}_{\mathbb{N}}$ into $\mathrm{LEV}_{\mathbb{N}}$ such that for any $x \in \mathrm{VAR}_{\mathbb{N}}$ and $n \in \mathbb{N}$, if $x \in \mathrm{VAR}_n$ then $A(x) \in \mathrm{LEV}_n$. Let $f \in \mathrm{SF}_n$ ($n \in \mathbb{N}^+$) and let $A$ be an assignment. The *value* of $f$ at $A$ is defined as follows. For some $g \in \mathrm{UF}_n$ and $a_1$, …, $a_m$ ($m \in \mathbb{N}^+$), $f = [\![g, a_1, \ldots, a_m]\!]$. For each $i$ with $1 \leq i \leq m$, define $b_i$ as follows: if $a_i \in \mathrm{LEV}_{\mathbb{N}}$, then $b_i = a_i$; if $a_i \in \mathrm{VAR}_{\mathbb{N}}$, then $b_i = A(a_i)$; if $a_i$ is $[\![\circ, h]\!]$ for some $h \in \mathrm{SF}_n$, then $b_i$ is the *value* of $h$ at $A$. The *value* of $f$ at $A$ is then $g(b_1, \ldots, b_m)$.

It is easy to show that all the assumptions and claims made in §4 hold in the model, when interpreted in the obvious way. These fall essentially into two groups: existence and distinctness claims. An example of an existence claim is as follows. In §4.2 I claimed that if $f$ is an $n$-place unstructured function of level 1, and each of $a_1$, …, $a_n$ is either an object, a level 0 variable, or $\circ S$, where $S$ is a structured function of level 1, then $f\langle a_1, \ldots, a_n\rangle$ is a structured function of level 1. This holds in our model because for any $n$-place $f' \in \mathrm{UF}_1$, and $a'_1$, …, $a'_n \in \mathrm{LEV}_0 \cup \mathrm{VAR}_0 \cup \{[\![\circ, g]\!] : g \in \mathrm{SF}_1\}$, $[\![f', a'_1, \ldots, a'_n]\!] \in \mathrm{SF}_1$ (by the definition of $\mathrm{SF}_1$). Similarly, the existence of unstructured functions such as $\neg$, $\exists_x$ (for $x$ a level 0 variable), and the raising functions of §4.6 follows easily from the axioms of ZFU. An example of a distinctness claim is the assumption made in §4.2 that no structured function is an unstructured function. This holds because for any $n \in \mathbb{N}^+$ and $f \in \mathrm{SF}_n$, $f$ is an ordered $m$-tuple (for some $m \in \mathbb{N}^+$). It follows that $f$ is finite. In contrast, for any $r \in \mathbb{N}^+$ and $g \in \mathrm{UF}_r$, $g$ is infinite (because $g$ is a set-theoretic function from $\mathrm{LEV}_{r-1}$ into $\mathrm{LEV}_{r-1}$, and $\mathrm{LEV}_{r-1}$ is infinite). Similarly for the other claims made in §4. It follows that the account of §4 is consistent.

## References


Anderson, C. A. 1989. Russellian Intensional Logic. In J. Almog, J. Perry and H. Wettstein (eds), *Themes from Kaplan*: 67–103. New York: Oxford University Press.

Chihara, C. S. 1973. *Ontology and the Vicious-Circle Principle*. Ithaca: Cornell University Press.

Church, A. 1976. Comparison of Russell's Resolution of the Semantical Antinomies with that of Tarski. *Journal of Symbolic Logic* 41: 747–60.

Forster, T. E. 1995. *Set Theory with a Universal Set: Exploring an Untyped Universe*. Second edition. Oxford: Clarendon Press.

Frege, G. 1892. On Sense and Reference. In P. Geach and M. Black (eds), *Translations from the Philosophical Writings of Gottlob Frege*, second edition, 1960: 56–78. Oxford: Basil Blackwell.

Goldfarb, W. 1989. Russell's Reasons for Ramification. In C. W. Savage and C. A. Anderson (eds), *Reading Russell: Essays in Bertrand Russell's Metaphysics and Epistemology*: 24–40. Minneapolis: University of Minnesota Press.

Gödel, K. 1944. Russell's Mathematical Logic. In his *Collected Works, Volume II: Publications 1938–1974*, S. Feferman *et al.* (eds), 1990: 119–41. New York: Oxford University Press.

Kaplan, D. 1977. Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals. In J. Almog, J. Perry and H. Wettstein (eds), *Themes from Kaplan*, 1989: 481–563. New York: Oxford University Press.

Klement, K. C. 2004. Putting Form Before Function: Logical Grammar in Frege, Russell, and Wittgenstein. *Philosophers' Imprint* 4: 1–47.
—2010. The Functions of Russell's No Class Theory. *Review of Symbolic Logic* 3: 633–64.

Linsky. B. 1999. *Russell's Metaphysical Logic*. Stanford: CSLI Publications.

Maddy, P. 1983. Proper Classes. *Journal of Symbolic Logic* 48: 113–39.

Prior, A. N. 1971. *Objects of Thought*. P. T. Geach and A. J. P. Kenny (eds). Oxford: Clarendon Press.

Ramsey, F. P. 1925. The Foundations of Mathematics. In his *Philosophical Papers*, D. H. Mellor (ed.), 1990: 164–224. Cambridge: Cambridge University Press.

Russell, B. 1903. *The Principles of Mathematics*. Second edition. New York: Norton.
—1908. Mathematical Logic as Based on the Theory of Types. *American Journal of Mathematics* 30: 222–62.

Stalnaker. R. C. 1984. *Inquiry*. Cambridge, MA: MIT Press.

Whitehead, A. N. and B. Russell. 1927. *Principia Mathematica, Volume I*. Second edition. Cambridge: Cambridge University Press.