

# Gödel vis-à-vis Russell: Logic and Set Theory to Philosophy

Juliet Floyd and Akihiro Kanamori  
Boston University

Gödel's work from the beginning to his first substantive explorations in philosophy would to a significant extent be contextualized by, reactive to, and reflective of, Russell's. Russell was the towering figure who set the stage for analytic philosophy in the first two decades of the 20<sup>th</sup> Century; Gödel insisted that his mathematical work was substantively motivated by his own philosophical outlook; and so it becomes especially pertinent to draw out and highlight the interconnections between the two.

What follows is a narrative that focuses on the interplay of several arching motifs: Russell's theory of types and Axiom of Reducibility; definability as analysis; Gödel's incompleteness theorems and constructible sets; and Russell's and Gödel's respective construals of the nature of truth as a, if not *the*, philosophical problem. More specifically, Gödel's reflections on Russell's so-called "multiple relation theory of judgment" (MRTJ), the theory of truth at work in the *Principia Mathematica* and later writings of Russell, especially his William James Lectures (1940), will be set forth and framed. This will allow us to draw out, in a provisional way, some of the philosophical significance of Gödel's recently transcribed *Max Phil* notebooks.

Russell's MRTJ was an idiosyncratic version of the correspondence theory of truth, framed in terms of complex (non-dual) relations between judgers, the objects of their judgments, and their environment. Our view is that it was this *specific* theory of truth that preoccupied Gödel by 1929, and again in 1942-1943, when he returned to reading Russell in earnest.

We offer a narrative of influence, reaction, and confluence, of suggested thematic causalities discernible in Gödel's intellectual development. With the conviction that there is still much to be said about Gödel vis-à-vis Russell, we forge a particular path through now well-known aspects of Gödel's work and thought, putting new weight on some lesser-known aspects. In particular, we give a nuanced reading of Gödel's mathematical platonism as emergent according to mathematical advances and, by his lights, increasingly indispensable. We also insist, especially given the evidence of the *Max Phil* notebooks, on the increasingly serious importance to Gödel (as also evidently to Russell throughout) of philosophical questions of the most classical and fundamental kind concerning the relation of logic to grammar and to mathematics, of truth to perception, and these questions' relevance to foundational questions about the very possibility of conceptual analysis. Others have weighed in here, in an increasingly informative and fine-grained way; our aim is primarily to broach the bearing of the

notebooks *Max Phil IX* and *X*, written during a period of engagement with Russell's philosophy, on Gödel's philosophical sensibility. While Gödel would later turn to serious reconsideration of Leibniz (whom he had read while a student in Vienna) and Husserl, his formative engagements were with Russellian moves, these ultimately forming the base camp for the later philosophical expeditions.

§1 casts light from our vantage point on Gödel's work in mathematical logic, with emphasis on truth, and §2, on his subsequent work in set theory, with emphasis on definability and an emergent mathematical platonism. These sections draw somewhat on Kanamori (2007), which serves as a general reference for details on Gödel's results and their historical context, here supplemented with sharpened emphasis on the arc of Gödel's engagement with Russell. §3 turns to philosophy, and an overview of the remainder of the paper: a reconstruction of Gödel's engagement with Russell's purely philosophical writings on truth, especially his "multiple relation theory of judgment" (MRTJ). §4 elaborates on the MRTJ, including Russell's treatment of "judgments of perception" and his writing on the theory of knowledge 1910-1913, keeping Gödel's appropriation of Russell constantly in view. §5 analyzes Gödel's reactions to Russell's *An Inquiry into Meaning and Truth* (1940) based on evidence emerging in *Max Phil IX-X*. §6 reviews sources and historical evidence for our claims.

## 1. Logic and truth

Gödel's salad days at the University of Vienna, where he turned from physics to mathematics and then to logic, have been given a rounded and balanced account;<sup>1</sup> it serves our purposes here to briefly draw out how he came to focus on logic and foundations in substantial part through early encounters with Russell. Below, in §6.1, we shall treat the issue of when Gödel may have come to Russell's purely philosophical writings. In his second year 1925-1926 at the university, when he was 19, Gödel participated in Moritz Schlick's seminar studying Russell's *Introduction to Mathematical Philosophy* (1919a) in German translation. This is likely to have been Gödel's first encounter with the Peano axioms, relations, cardinal and ordinal numbers, the Axiom of Choice, and propositional functions. Throughout, Gödel attended meetings of Schlick's circle, the Vienna Circle, where Russell was discussed and, by 1928, the turn to logical empiricism took place under the influence of Wittgenstein's *Tractatus*.<sup>2</sup> In the summer of 1928, Gödel purchased a copy of *Principia Mathematica*,

1 Cf. Dawson (1997, chap. II).

2 Olga Taussky-Todd (1987: p. 35) in remembrances of Gödel recalled his participation in Schlick's seminar; she wrote (p. 40) of the circle: "The Wiener Kreis, a successor of the Mach Verein, was a creation of Schlick. Time was ripe for such a creation in the late twenties. For, the necessity to test the foundation of mathematical thinking, the methods of proofs, the axioms, the rules, became pressing. It was a time of 'Sturm und Drang'. There were Peano's axioms of the integers, as they are described in Russell's 'Introduction' and the edifice of the *Principia Mathematica* above them. One found a way of explaining the 'Paradoxes'".

Much, of course, has been written about the Vienna Circle. See e.g. the well-documented Stadler (2001), according to which Gödel participated as early as winter 1924 (p. 206) and as late as summer 1931 (p. 297). Gödel, like Wittgenstein himself earlier, would soon disavow the circle's logical empiricism.

presumably the second edition (1927).<sup>3</sup> By then it was clear that Gödel would write his dissertation with Hans Hahn, who happened to be versed in *Principia*. In the winter semester of 1928-1929, Gödel evidently took Rudolf Carnap's course "The Philosophical Foundations of Arithmetic".<sup>4</sup>

With this assimilation of logic, Gödel quickly made prodigious progress. By the summer of 1929, he had established the Completeness Theorem (1930), the semantic completeness of first-order logic, for his doctoral dissertation. (1930) begins by setting the stage with *Principia Mathematica*, and the source of the completeness problem is the (1928) Hilbert-Ackermann monograph, whose *Principia*-inspired calculus is closely followed in (1930). By the summer of 1930, he had established the celebrated Incompleteness Theorem (1931), the syntactic incompleteness of formal arithmetic.

What was Gödel's way in? Carnap through his course has been accorded significant influence on Gödel's work, but this devolves, in the end, to providing language for discussing consistency and completeness.<sup>5</sup> Carnap's course was based on a book manuscript that he was then working on, the surviving first part of which has now been published posthumously.<sup>6</sup> One sees that the results there are routine or even tautological. Alfred Tarski lectured in Karl Menger's colloquium in February 1930 on metamathematics, but he put the emphasis on a separate language for expressing consistency and provability.<sup>7</sup> So what about Gödel's specific approach and techniques? For this there is provenance, if provenance is to be found, in *Principia Mathematica* itself and its engagement with truth and definability. To evidence this, we quickly recall the logic of the *Principia*, the ramified theory of types, and then quote some passages.

Russell's *ramified theory of types* is a scheme of logical definitions parameterized by *types* and *orders* indexed by the natural numbers. Propositions are classified according to the notion of *propositional function*, which when taking an argument (input) becomes a proposition. Propositional functions are incipiently *intensional*, in that what is a possible argument is not decided beforehand; for it to have an *extensional* counterpart is to have comprehended a class consisting exactly of the arguments that yield propositions. At the cost of some extensional distortion, the scaffolding structure can be succinctly conveyed by taking "object" for "propositional function" and "constituent" for an argument for which there is a corresponding proposition that holds. We may then say that the universe is to consist of *objects* stratified into disjoint types  $T_n$ , where  $T_0$  consists of the *individuals*,  $T_{n+1} \subseteq \{X \mid X \subseteq T_n\}$ , and the types  $T_n$  for  $n > 0$  are further ramified into the orders  $O_n^i$  with  $T_n = \bigcup_i O_n^i$ . An object in  $O_n^i$  is to be defined either in terms of individuals or of objects in some fixed  $O_m^j$  for some  $j < i$  and  $m \leq n$ , the definitions allowing for quantification over only the  $O_m^j$ . This precludes Russell's Paradox and other "vicious circles", as objects can consist only of previous objects and are built up through definitions referring only to previous stages. However, in this system it is

3 Dawson (1997: p. 53, n. 5). Citing book records, Dawson (p. 53f) dated Gödel's shift away from more classical mathematical fields toward logic and foundations as having occurred between the summer and fall of 1928.

4 Dawson (1997: p. 27). Awodey and Carus (2010: p. 254, n. 4).

5 Goldfarb (2005) and Awodey and Carus (2010).

6 Carnap (2000).

7 Dawson (1997: p. 62).

impossible to quantify over all objects in a type  $T_n$ , and this makes the formulation of numerous mathematical propositions at best cumbersome and at worst impossible. So Russell was led to introduce his *Axiom of Reducibility*, which asserts that *for each object there is a predicative object having exactly the same constituents*, where an object is *predicative* if its order is the least greater than that of its constituents. This axiom reduced consideration to individuals, predicative objects consisting of individuals, and so on—the *simple theory of types*.<sup>8</sup>

The ramification into orders accommodates the following passage, written by 1910 from *Principia*, Introduction, chapter II, Section III “Definition and Systematic Ambiguity of Truth and Falsehood”, p. 42ff:

That the words “true” and “false” have many different meanings, according to the kind of proposition to which they are applied, is not difficult to see. Let us take any function  $\varphi\hat{x}$  and let  $\varphi a$  be one of its values. Let us call the sort of truth which is applicable to  $\varphi a$  “*first truth*.” (This is not to assume that this would be first truth in another context; it is merely to indicate that it is the first sort of truth in our context.) Consider now the proposition  $(x) . \varphi x$ . If this has truth of the sort appropriate to it, that will mean that every value  $\varphi x$  has “*first truth*.” Thus if we call the sort of truth that is appropriate to  $(x) . \varphi x$  “*second truth*,” we may define “ $\{(x) . \varphi x\}$  has second truth” as meaning “every value for  $\varphi\hat{x}$  has first truth”, *i.e.* “ $(x) . \varphi x$  has first truth.” Similarly, if we denote by “ $(\exists x) . \varphi x$ ” the proposition “ $\varphi x$  sometimes,” *i.e.* as we may less accurately express it, “ $\varphi x$  with some value of  $x$ ,” we find that  $(\exists x) . \varphi x$  has second truth if there is an  $x$  with which  $\varphi x$  has first truth; thus we may define “ $\{(\exists x) . \varphi x\}$  has second truth” as meaning “some value for  $\varphi\hat{x}$  has first truth,” *i.e.* “ $(\exists x) . (\varphi x \text{ has first truth})$ .” Similar remarks apply to falsehood. Thus “ $\{(x) . \varphi x\}$  has second falsehood” will mean “some value for  $\varphi\hat{x}$  has first falsehood,” *i.e.* “ $(\exists x) . (\varphi x \text{ has first falsehood})$ ,” while “ $\{(\exists x) . \varphi x\}$  has second falsehood” will mean “all values for  $\varphi\hat{x}$  have first falsehood,” *i.e.* “ $(x) . (\varphi x \text{ has first falsehood})$ .” Thus the sort of falsehood that can belong to a general proposition is different from the sort that can belong to a particular proposition.

Seen against a longstanding tradition of truth as unanalyzable and primordial, this is actually a striking statement of variegation about meaning and sense.<sup>9</sup> Set out here is a notably early recursive definition, the clauses appearing to us latterly as reminiscent of a Tarskian definition of satisfiability. However, there is a crucial difference in that, with no external vantage point, “has second truth” and the like are being woven *into* the

8 The second, 1927 edition of *Principia* differs from the first edition mainly in having a further introduction and three appendices, mainly devoted to the possible reach and changes one needed if one were to eschew, as suggested by Wittgenstein in the *Tractatus*, the Axiom of Reducibility. Gödel (1944: p. 145f) in fact found a flaw in one of the new lemmas, \*89.16, and considered the issue that it was addressing to remain open. Landini (2007) has worked out a syntax for the second edition of the *Principia* and shown (p. 211) how to get around \*89.16; Hazen and Davoren (2000) and Hazen (2000) also reconstruct the ramified type theory substitutionally, offering some corrections. In any case, Gödel in his salad days would have seen the workings of the Axiom of Reducibility in the *Principia* text and subsequently put much store in it, as discussed in the next section.

9 In Frege’s *Grundgesetze* (1893, §§22–25), propositions have signification [*Bedeutung*], either the True [*das Wahre*] or the False [*das Falsche*], and this ascription is evidentially compositional. However, truth itself is not split, and there would famously be inconsistency.

language through definition and deft use of quotation marks.<sup>10</sup> *Principia* then proceeds to the basis for the recursion, “a definition of the simplest kind of *truth* and *falsehood*”. This *Ur*-arena, the foundational beginning of the hierarchy, is just where more purely philosophical issues about truth enter in. We shall explicate these passages in §§3-5, for they were crucial stepping stones for Gödel’s entry into philosophy.

Soon afterwards in *Principia*, the following passage occurs in Introduction, chapter II, section VIII “The Contradictions”, p. 61:

Richard’s paradox ... is akin to that of the least definable ordinal. It is as follows: Consider all decimals that can be defined by means of a finite number of words; let  $E$  be the class of such decimals. Then  $E$  has  $\aleph_0$  terms; hence its members can be ordered as the 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, ... Let  $N$  be a number defined as follows: If the  $n$ th figure in  $n$ th decimal is  $p$ , let the  $n$ th figure in  $N$  be  $p + 1$  (or 0, if  $p = 9$ ). Then  $N$  is different from all members of  $E$ , since, whatever finite value  $n$  may have, the  $n$ th figure in  $N$  is different from the  $n$ th figure in the  $n$ th of the decimals composing  $E$ , and therefore  $N$  is different from the  $n$ th decimal. Nevertheless we have defined  $N$  in a finite number of words, and therefore  $N$  ought to be a member of  $E$ . Thus  $N$  both is and is not a member of  $E$ .

At the cusp of the logical and the linguistic, Richard’s Paradox is a particular, structured paradox about definability in relation to truth, to be located at the crossing of intensionality and extensionality. The negotiation of this intersection would remain, as we shall see, a crucial challenge both for Russell and for Gödel, and philosophical commonalities in their approach to it—in contrast to Carnap’s and Tarski’s—would be emphasized by Gödel, even more than Russell, over the following decade and a half.

Getting to Gödel, he later maintained that he had come to the indefinability of truth for arithmetic in the summer of 1930, *before* getting to incompleteness.<sup>11</sup> Gödel

10 According to Landini (1998: p. 282f), Russell had abandoned his former ontology of propositions, and: “With a no-propositions theory, ‘truth’ and ‘falsehood’ no longer need be regarded as primitive properties. Rather, there are now to be different senses of ‘truth’ and ‘falsehood’ as applied to statements differing in structure. The different senses, in turn, explain and philosophically justify the order part of order/type indices on a predicate variable”. Landini went as far as to write (p. 291): “Once we see that it was a recursive truth definition that generates orders, we get an entirely new perspective on *Principia*”. Given our focus on Gödel, Russell, and the nature of truth however, we shall draw out threads of both philosophers’ thinking that join them more closely at the hip.

11 Gödel remarked, later in 1976, as conveyed by (Wang 1996: p. 82):

In summer 1930 I began to study the consistency problem of classical analysis. It is mysterious why Hilbert wanted to prove directly the consistency of analysis by finitary methods. I saw two distinguishable problems: to prove the consistency of number theory by finitary number theory and to prove the consistency of analysis by number theory. By dividing the difficulties, each part can be overcome more easily. Since the domain of finitary number theory was not well defined, I began by tackling the second half: to prove the consistency of analysis relative to full number theory. It is easier to prove the *relative* consistency of analysis. Then one only has to prove by finitary methods the consistency of number theory. But for the former one has to assume number theory to be true (not just the consistency of a formal system for it).

I represented real numbers by predicates in number theory [which express properties of natural numbers] and found that I had to use the concept of truth [for number theory] to verify the axioms of analysis. By the enumeration of symbols, sentences, and proofs of the given system, I quickly discovered that the concept of arithmetic truth cannot be defined in arithmetic.

would have established this, as Tarski (1935, §5, theorem I) did later for general structures crediting Gödel (1931), with a simple version of the later incompleteness argument: Enumerate the formulas in one free variable and derive a contradiction from the definability of the predication that the  $n$ th formula with the numeral  $n$  is true. One can press here the thematic connections with *Principia*: there is first of all a *formula-by-formula working out of truth*; there is presumption that a *general definition is ambiguous* in application; and there is *the operational appeal* to Richard's Paradox to preclude definability. Even more, questions about what Russell called (1908: p. 395, 1910b: ch. V ¶5) "*the fundamental question [of philosophy], namely, the nature of truth*", emerge quite explicitly in *Principia* at just this point, permeating Russell's presentation of types. This did not escape Gödel's eye, as we shall see.

The central idea of the Incompleteness Theorem is the formal definability of provability, which through the arithmetization of syntax leads to an unprovable formula based on the coded predication that  $n$ th formula in one free variable with the numeral  $n$  is provable. This subsumption of metamathematics into mathematics is unquestionably a major conceptual breakthrough, but in the pilgrim's progress there are signposts in light of *Principia*. Russellian recursion for formulas is worked out, though not for orders of truth but for provability, which thus *can* be recursively woven into the language, and Richard's Paradox is specifically implemented to transcend definability. As Gödel wrote (1931, p. 175): "The analogy of this argument with the Richard antinomy leaps to the eye".

If Russellian themes can be discerned in Gödel (1931), so also a parting of the ways, one having to do with the way in which mathematical truth as such is framed. Although "truth" does not appear in (1931), Gödel was emphatic about the difference between truth and provability in terms of definability at a 2 July 1931 meeting of the Vienna Circle<sup>12</sup> and in a 12 October 1931 letter to Ernst Zermelo.<sup>13</sup> Retrospectively, Gödel in a 7 December 1967 letter to Hao Wang was forceful about how truth contrasts with and frames provability.<sup>14</sup>

---

Gödel had written earlier in an unsent 1970 letter to one Yossef Balas (Gödel 2003a: p. 9ff):

The occasion for comparing truth and demonstrability was an attempt to give a relative model-theoretic consistency proof of analysis in arithmetic. This leads almost by necessity to such a comparison. (For, an arithmetical model of analysis is nothing else but an arithmetical  $\epsilon$ -relation satisfying the comprehension axiom):  $(\exists n) (x) [x \in n \equiv \varphi(x)]$ .

- 12 Stadler (2001: p. 297f) records Gödel at a discussion of Carnap's metalogic:

Well, it has been proven that certain metamathematical concepts are definable in the same language, e.g., 'formula', 'provable formula', and in general all the concepts which, e.g., Hilbert uses. By contrast, there are other metamathematical concepts, which cannot be defined in the same language; i.e., if one were to use them, it would lead to a circle antinomy. One example for this in your system is the concept 'correct [*richtige*] numerical formula'.

- 13 See Gödel (2003b: p. 423ff). In the letter, Gödel pointed out a quick proof of incompleteness: The class of provable formulas is definable and the class of true formulas is not, and so there must be a true but unprovable formula.

- 14 See Gödel (2003b: p. 396ff). Gödel wrote:

...the heuristic principle of my construction of undecidable number theoretical propositions in the formal systems of mathematics is the highly transfinite concept of 'objective mathematical truth', as opposed to that of 'demonstrability', with which it was generally confused before my own and Tarski's work. Again the use of this transfinite concept eventually leads to finitarily provable results, e.g., the general theorems about the existence of undecidable propositions in consistent formal systems.

Gödel's (1931), submitted for publication 17 November 1930, had *Principia* in the title and was written in its thematic wake, with the actual formalization carried out in a system  $P$  which was essentially the logic of *Principia* superposed on the Peano axioms with the natural numbers as individuals. The paper had a prescient footnote 48a, which signaled a further move beyond *Principia* with respect to truth:

As will be shown in Part II of this paper, the true reason for the incompleteness inherent in all formal systems of mathematics is that the transformation of ever higher types can be continued into the transfinite (cf. D. Hilbert, "Über das Unendliche", Math. Ann. 95, p. 184), while any formal system at most denumerably many of them are available. For it can be shown that the undecidable propositions constructed here become decidable whenever appropriate higher types are added (for example, the type  $\omega$  to the system  $P$ ). An analogous situation prevails for the axiom system of set theory.

This footnote has been made much of.<sup>15</sup> The "48a" would suggest a late addition, and Gödel had read Hilbert (1926) (only) in 1930.<sup>16</sup> Hilbert's paper featured an approach to establishing the Continuum Hypothesis through transfinite recursion, and while not successful, it stimulated Gödel to entertain transfinite types. Having transfinite types dramatically changes the tenor of *Principia*, in that types there had only been providing a background logical structure, whereas now types were on their way to becoming classes of the system in the explicit indexing.<sup>17</sup> The last sentence of the footnote calls to mind the other reference to set theory in the paper; Gödel wrote (1931, p. 178) of his comprehension axiom IV, foreshadowing his approach to set theory, "This axiom plays the role of the axiom of reducibility (the comprehension axiom of set theory)".

The following appeared in a summary (1932), dated 22 January 1931, of a talk on the incompleteness results given in Menger's colloquium. Matters in a footnote, perhaps an afterthought then, have now been expanded to take up fully one-third of an abstract on incompleteness (1990, p. 237).

If we imagine that the system  $Z$  [first-order Peano arithmetic] is successively enlarged by the introduction of variables for classes of numbers, classes of classes of numbers, and so forth, together with the corresponding comprehension axioms, we obtain a sequence (continuable into the transfinite) of formal systems that satisfy the assumptions mentioned above, and it turns out that the consistency ( $\omega$ -consistency) of any of those systems is provable in all subsequent systems. Also, the undecidable propositions constructed for the proof of Theorem 1 [the Gödelian sentences] become decidable by the adjunction of higher types and the corresponding axioms; however, in the higher systems we can construct other undecidable propositions by the same procedure, and so forth. To be sure, all the propositions thus constructed are expressible in  $Z$  (hence are number-theoretic propositions); they are, however, not decidable in  $Z$ , but only in higher systems for example, in that of analysis. In case we adopt a type-free construction of mathematics, as is done in the axiom system of set theory, axioms of cardinality (that is, axioms postulating the existence of sets of ever higher cardinality) take the place of

15 See e.g. Kreisel's memoir on Gödel (1980: p. 183, 195, 197) and Feferman (1987), where the view advanced in the footnote is referred to as "Gödel's doctrine".

16 Cf. Wang (1996: p. 73).

17 *Principia*, p. 53:

We do not arrive at functions of an infinite order, because the number of arguments and of apparent variables in a function must be finite, and therefore every function must be of a finite order. Since the orders of functions are only defined step by step, there can be no process of 'proceeding to the limit,' and functions of an infinite order cannot occur.

type extensions, and it follows that certain arithmetic propositions that are undecidable in  $Z$  become decidable by axioms of cardinality, for example, by the axiom that there exist sets whose cardinality is greater than every  $\alpha_n$ , where  $\alpha_0 = \aleph_0$ ,  $\alpha_{n+1} = 2^{\alpha_n}$ .

This passage can be considered the point of transition in Gödel's progress from type theory to set theory. There are ever more types, here domains of "systems". Although "truth" is again not explicit, truth figures centrally through the "comprehension axioms". Gödel pointed out in a letter to Paul Bernays of 2 April 1931 and in letters to Carnap of 11 September and 28 November 1932 how truth for a given type *can* be defined in the next higher type, where quantification of *arbitrary* subsets of the given type is possible, and, with a truth predicate available, how previously undecidable propositions become newly decidable.<sup>18</sup> In the passage Gödel is specific, with a nod to his "Second Incompleteness Theorem" on consistency, about the consistency of a system being newly provable in subsequent systems and the various undecidable propositions being number-theoretic, expressible already at the bottom level  $Z$ .

This hierarchical development as a pursuit of truth contrasts in thrust and structure to the approach of Tarski (1935), with its set-theoretic definition of satisfaction-in-a-structure and language vs. metalanguage distinction. Having apprehended the indefinability of truth for a type *at* the type, Gödel sees its definability specifically in the *next* type as metalanguage "system", with the *formalization* of satisfaction there through "comprehension axioms".<sup>19</sup> The process is iterated, generating ever more newly decidable number-theoretic propositions. Notably, one can see Gödel's approach as more consonant with Russell's proposal for a hierarchy of languages to get at truth in his 1922 introduction to Wittgenstein's *Tractatus* (cf. §4.4 below).

The last sentence of the passage is Gödel's first remark of substance on set theory, and it foreshadows his engagement with the Axiom of Replacement (cf. §2), that for any (definable) class function  $F$  and set  $a$ , the image  $F "a = \{F(x) \mid x \in a\}$  is a set. Gödel's example of an "axiom of cardinality" evidently has to do with the successive cardinalities of the first infinite types as given by taking power sets and, notably, is the one that both Abraham Fraenkel and Thoralf Skolem in the early 1920s had pointed to as unprovable in Zermelo's 1908 axiomization of set theory and used to motivated Replacement.<sup>20</sup>

## 2. Set theory and definability

Gödel delivered an invited lecture titled "The present situation in the foundations of mathematics" on 30 December 1933 at the annual meeting of the Mathematical Association of America, and the incisive text (1933o) of the lecture illuminates his further thinking at the time about the motifs exhibited in the quoted passages above and their projection into set theory. Gödel initially set out the problem of giving a foundation for mathematics as falling into two parts, reducing the methods of proof to a minimum of axioms and rules of inference and then justifying these.

<sup>18</sup> Cf. Gödel (2003a: pp. 97ff, 347ff, 355ff).

<sup>19</sup> Tarski (1935: p. 194, n. 108) acknowledged Gödel's (1931) footnote 48a.

<sup>20</sup> See for example Kanamori (2012). With the natural numbers being in the system as forming a set, so then would  $\{\alpha_0, \alpha_1, \alpha_2, \dots\}$  by the axiom.

Gödel considered the first part to have been satisfactorily solved by the “formalization” of mathematics, which can be carried out in a set-theoretic version of Russell’s simple theory of types. Gödel in fact viewed (p. 3f) “the system of axioms for the theory of aggregates, as presented by Zermelo, Fraenkel, and von Neumann” as “nothing else but a natural generalization of the theory of types, or rather, what becomes of the theory of types if certain superfluous restrictions are removed”. Specifically crediting von Neumann (1929), Gödel described these, of which there are mainly two. First, instead of having disjoint types with sets of type  $n+1$  consisting purely of sets of type  $n$ , sets can be *cumulative* in the sense that sets of type  $n$  can consist of sets of *all* lower types. If  $S_n$  is the collection of sets of type  $n$ , then:  $S_0$  is the type of the individuals, and recursively,  $S_{n+1} = S_n \cup \{X \mid X \subseteq S_n\}$ . Second, the process can be continued into the transfinite, starting with the cumulation  $S_\omega = \bigcup_n S_n$ , proceeding through successor stages as before, and taking further unions at limit stages. Gödel here had broached the cumulative hierarchy picture of sets, what would come to be called the *iterative conception of set*.

In the axiomatic development of set theory, Zermelo (1908) had axiomatized “the set theory of Cantor and Dedekind” partly for rigorizing his well-ordering arguments. Work most substantially of von Neumann in the 1920s led to the incorporation of Cantor’s ordinal numbers as *bono fide* sets, the ordinals—transitive sets well-ordered by membership—and the Axiom of Replacement for the formalization of transfinite recursion. In his last axiomatization paper (1929), von Neumann formulated the Axiom of Foundation, that every set is well-founded, and defined the *cumulative hierarchy* in his system via transfinite recursion. The axiom entails, in modern terms, that the universe  $V$  of sets is globally structured through a stratification into cumulative “ranks”  $V_\alpha$ , where:  $V_0 = \emptyset$ ;  $V_{\alpha+1} = P(V_\alpha)$ , the power set;  $V_\delta = \bigcup_{\alpha < \delta} V_\alpha$  for limit ordinals  $\delta$ ; and  $V = \bigcup_\alpha V_\alpha$ . Zermelo in his remarkable (1930) subsequently provided his final axiomatization of set theory, incorporating both Replacement and Foundation, and through these axioms established a second-order categoricity of sorts for set theory with the cumulative hierarchy picture.

While this development to the cumulative hierarchy picture took decades, Gödel had come to the picture forthwith and informally from the simple theory of types as a matter of defining truth in successive transfinite extensions. In later years, the cumulative hierarchy picture would come to be seen as *the* way to view the set-theoretic universe, and in a notable inversion, the iterative conception of set would be latterly invoked to *motivate* the axioms of set theory.

Proceeding with his lecture (1933o), Gödel pointed out (p. 8):

... in order to state the axioms for a formal system, including all the types [indexed] up to a given ordinal  $\alpha$ , the notion of this ordinal  $\alpha$  has to be presupposed to be known, because it will appear explicitly in the axioms. On the other hand, a satisfactory definition of transfinite ordinals can be obtained only in terms of the very system whose axioms are to be set up. I don’t think that this objection is serious for the following reason: The first two or three [infinite] types already suffice to define very large ordinals. So you can begin by setting up axioms for these first types, for which purpose no ordinal whatsoever is needed, then define a transfinite ordinal  $\alpha$  in terms of these first few types and by means of it state the axioms for the system, including all classes of type less than  $\alpha$ .

This is a penetrating statement about the possibilities for generating higher types, with truth being implicitly brought into play in a first look at the definability of the

indexing ordinals themselves. To get “the types [indexed] up a given ordinal  $\alpha$ , the notion of this ordinal  $\alpha$  has to be presupposed to be known”, but this is not a “serious” objection as “the first two or three [infinite] types already suffice to define very large ordinals” with which one can then define and index higher types. The “ordinals” here would actually be well-orderings, not (von Neumann) ordinals; how one codes the indexing in such bootstrapping “autonomous progressions”, in later terminology, evidently becomes an issue.<sup>21</sup>

Gödel turned forthwith (p. 9a) to the role of the axioms of set theory, and his discussion here reveals an equivocation. Gödel wrote that the axioms of set theory “can be characterized by a certain closure property” on the cumulative hierarchy which for limit stages says “if  $M$  is a set of ordinals definable in the system and if to each ordinal of  $M$  you assign a type contained in the system, then the type obtained by summing up those types is also in the system”. This is tantamount to the Axiom of Replacement. But later, Gödel wrote that “the totality of all systems thus obtained seems to form a totality of similar character to the set of ordinals of the second ordinal class”. However, the Cantorian second class, of the countable ordinals, is but a small initial segment of the ordinals as given by Replacement. This equivocation fits Gödel’s grappling with how far to go in a refined, definability setting (see below).

In the second part of his lecture Gödel turned to the justification of axioms and rules of inference, and here he found the situation “extremely unsatisfactory”, opining that “as soon as we come to attach a meaning to our symbols serious difficulties arise”. In setting out these, Gödel dwelled on non-constructive existence proofs and more so on impredicative definitions, as in defining a property  $P$  by (p. 18f): “An integer  $x$  shall possess the property  $P$  if for all properties (including  $P$  itself) some statement about  $x$  is true”.

. . . this process of definition presupposes that the totality of all properties exists somehow independently of our knowledge and definitions, and that our definitions merely serve to pick out certain of these previously existing properties. If we assume this, the method of non-predicative definition is perfectly all right . . .

But the situation becomes entirely different if we regard the properties as *generated* by our definitions. For it is certainly a vicious circle to generate an object by reference to a totality in which this very object is supposed to be present already.

Note the resonance here generally with Russell’s grappling with types and particularly with the previously quoted passage, in which “ $\alpha$  has to be presupposed as known” and autonomous progression is the way out.

Gödel soon continued (p. 19):

The result of the proceeding discussion is that our axioms, if interpreted as meaningful statements, necessarily presuppose a kind of Platonism, which cannot satisfy any critical mind and which does not even produce the conviction that they are consistent.

Nevertheless, Gödel found the prospects of establishing the *consistency* of the axioms promising e.g. “as the statement to be proved is of a very simple character and does not involve any of those objectionable notions such as ‘property of integers’”. Of course, such a proof “must be conducted by perfectly unobjectionable methods, i.e. it must

21 Gödel had arrived at the idea of autonomous progressions already when writing (1931), and it has led to important ramifications about finitary mathematics; see van Atten (2014).

strictly avoid the non-constructive existence proofs, non-predicative definitions and similar things, for it is exactly a justification of these doubtful methods which we are seeking”.

Gödel concluded the lecture with extended speculations about prospects for carrying out such a relative consistency proof in terms of an extension of intuitionistic mathematics, discussing Hilbert’s foundational program and the possibilities for proceeding from an intuitionistic standpoint.

The above quoted passage has become singled out as an outlier in light of Gödel’s retrospective pronouncements in his later years about how he had espoused platonism from the beginning and how important it was for the success of his work.<sup>22</sup> Note however that the passage is not inconsistent with the tone of the lecture, e.g. with the “exists somehow independently of our knowledge” of the previously quoted passage. In any case, one can take the lecture as a gauge of Gödel’s upward progress and changing attitude.

It has latterly been said that Gödel’s much-vaunted platonism may have actually been a development of growing conviction.<sup>23</sup> We argue for this on the *specific* grounds of Gödel’s progress in getting at higher and higher truth in steady, expansive reaction to Russell. Through the succession of the quoted (1931) footnote 48a, the 1931 abstract, and the 1933 lecture, one can see a pilgrim’s progress of gradual embrace. There is first the need for truth to frame unprovability for arithmetic; then the securing of truth in the “next higher system”; the self-fueling need for higher and higher systems; and thus the entrée into full-blown set theory with explicit indexing given by ordinals in autonomous progression. At work is a sort of progressive, step-by-step indispensability argument for mathematical platonism. With each successive success, more and more objectification becomes indispensable. And soon, Gödel would delimit his (1933o) reservations about impredicativity through a refined analysis.

Gödel, according to Wang (1981: p. 129) reporting on conversations with him in 1976, had already been working on Cantor’s continuum problem for some time and had devised what he considered to be a transfinite extension of the *ramified* theory of types. (One presumes some inspiration drawn here from Hilbert (1926), this time with its attempted proof of the Continuum Hypothesis through a “ramification” of recursive definitions.) Gödel in effect had started working up the *constructible hierarchy*, where in modern terms  $\text{def}(x) = \{y \subseteq x \mid y \text{ is first-order definable over } \langle x, \in \rangle\}$ ,<sup>24</sup> and:  $L_0 = \emptyset$ ,  $L_{\alpha+1} = \text{def}(L_\alpha)$ , and  $L_\delta = \bigcup_{\alpha < \delta} L_\alpha$  for limit ordinals  $\delta$ . With satisfaction or truth, for the

22 See in particular Gödel’s letter of 7 December 1957 to Hao Wang Gödel (2003b: p. 396ff). Gödel (2003a: p. 446ff) in 1974 wrote in a draft letter in response to a questionnaire, “I was a mathematical and conceptual realist since about 1925”. Gödel was hesitant in English in the early 1930s and would have people check his writing. John Dawson has pointed out that in the above quoted passage Gödel may have intended a sense better carried without the comma after “Platonism”; indeed, a further change of the succeeding “which” to “that” would have significantly softened the sense.

23 See especially Davis (2005: §2), and also Parsons (1995: §2), Goldfarb (2005: p. 249).

24 To elaborate,  $y$  is first-order definable over  $\langle x, \in \rangle$  if there is a first-order set-theoretic formula  $\varphi(v_0, v_1, \dots, v_n)$  and  $b_1, \dots, b_n \in x$  such that  $y = \{a \in x \mid \langle x, \in \rangle \models \varphi[a, b_1, \dots, b_n]\}$  with  $\models$  the satisfaction predicate.

structure  $\langle x, \in \rangle$ , needed to get at  $\text{def}(x)$ , Gödel's pursuit of *truth* in higher and higher systems was newly transmuted into a pursuit of hierarchical *definability*.<sup>25</sup>

There is evidence however that Gödel could not get far pursuing his envisioned (1933o) autonomous progression. Wang (1981: p. 129) reported how in Gödel's efforts to go up the constructible hierarchy he "spoke of experimenting with more and more complex constructions [of ordinals for indexing] for some extended period somewhere between 1930 and 1935". More pointedly, Kreisel in his memoir wrote (1980: p. 193): "As early as 1931, Gödel alluded to some reservations [about Replacement]". Kreisel continued (p. 196, with  $C_{\omega+\omega}$  his notation for  $V_{\omega+\omega}$ ): "In keeping with his reservations, mentioned on p. 193, Gödel first tried to do without the replacement property, and to describe the constructible hierarchy  $L_\alpha$  only for  $\alpha < \text{card } C_{\omega+\omega}$ ; in particular, without using von Neumann's canonical well-orderings. Instead, well-orderings had to be defined (painfully) in  $C_{\omega+\omega}$ ...". This is consistent with Gödel's equivocation mentioned earlier involving Replacement; Gödel was trying to do just with well-orderings, and only up to those of the cardinality described at the end of his abstract (1932).

The full embrace of Replacement led to a dramatic development. Set theory reached a new plateau with Gödel's formulation of the class  $L = \bigcup \{L_\alpha \mid \alpha \text{ is an ordinal}\}$  of *constructible* sets with which he established the relative consistency of the Axiom of Choice in mid-1935 and the Generalized Continuum Hypothesis in mid-1937. Gödel had continued the indexing of his hierarchy through *all* the ordinals as given beforehand to get a class model  $L$  of set theory and thereby to achieve *relative* consistency results, by showing that  $L$  satisfies the Axiom of Choice and the Continuum Hypothesis. His earlier (1933o) idea of using large ordinals (really, well-orderings) defined in low types in a bootstrapping process would not suffice. With the new focus on definability beyond truth, the (von Neumann) ordinals would be the spine for a definability hierarchy of sets, and this would be the key to the relative consistency results.

In his first announcement of  $L$ , (1938) Gödel conspicuously noted Russellian connections, which he would subsequently elaborate:

[The] "constructible sets" are defined to be those sets which can be obtained by Russell's ramified hierarchy of types, if extended to include transfinite orders. The extension to transfinite orders has the consequence that the model satisfies the impredicative axioms of set theory, because an axiom of reducibility can be proved for sufficiently high orders.

Gödel thus saw his new accomplishment as a working out of the ramified theory of types with the Axiom of Reducibility, newly implementable because of transfinite orders!

25 Gödel was thus working with satisfaction-in-a-structure concurrently as Tarski (1935) was formulating this notion for general structures. For Tarski too definability was and would be a central concern. In his (1931), written at the same time as the Polish original of his (1935), Tarski gave a mathematically precise (that is, set-theoretic) formulation of the informal concept of a (first-order) definable set of reals.

In the succeeding account (1939a), Gödel summarily presented the  $L_\alpha$  hierarchy in terms of “propositional functions”.<sup>26</sup> Gödel then sketched a proof of the Generalized Continuum Hypothesis in  $L$  and simply argued for the ZFC axioms holding in  $L$  as evident from the construction. Only toward the end did he mention (p. 31) that  $L$  “can be defined and its theory developed in the formal systems of set theory themselves”. However, this is a crucial point without which *relative* consistency cannot be secured. In fact, requisite is to show  $L^L = L$ , i.e.  $L$  constructed as a class in  $L$  is again  $L$ , for which *absoluteness* arguments, such as for the definition of (von Neumann) ordinals giving extensionally the same sets in different contexts, are needed.

In his classic monograph (1940), based on 1938 lectures, Gödel duly established  $L^L = L$ . He did this, as well as establish his relative consistency results, through a generation of  $L$  set by set by transfinite recursion with eight set-building operations. Thus having considerably distanced the context of discovery from the context of justification, Gödel rather obscured the proof of the Continuum Hypothesis in  $L$  but on the other hand crystallized how  $L$  can be seen to develop through a veritable “Gödel numbering” based on definitions and the extent of the ordinals.

However formal his published accounts, Gödel delivered two lucid lectures (1939b) and (1940a) on  $L$ , which brought out how he conceived of his work in continuing engagement with and reaction to Russell. In (1939b), notably delivered on 15 December 1939 at Hilbert’s Göttingen, Gödel (p. 6f) tellingly motivated the constructible sets through the ramified theory of types, with the successive orders of that theory to be given by hierarchical definition. Thus independently presented, he then described how quantification over higher types can be “intermixed”. With this preamble, Gödel (p. 8f) described “the objects of which set theory speaks” as falling into a “transfinite sequence of Russellian types”, the cumulative hierarchy of his (1933o). And then Gödel finally formulated the constructible sets as the  $L_\alpha$  hierarchy, shifting in effect the focus from Russellian types to orders. Gödel tellingly defined the “order” of an  $x \in L$  as that  $\alpha$  such that  $x \in L_{\alpha+1} - L_\alpha$ .

Gödel went on (p. 13) to describe the now well-known “Skolem hull” lemma that he had proved in (1939a) for his proof of the Continuum Hypothesis in  $L$ , asserting that “this fundamental theorem constitutes the corrected core of the so-called Russellian axiom of reducibility”. He continued (p. 14f):

This character of the fundamental theorem as an axiom of reducibility is also the reason why *the axioms of classical mathematics* hold for the model of the constructible sets. For after all, as Russell showed, the axioms of reducibility, infinity and choice are the only axioms of classical mathematics that do not have [a] tautological character. To be sure, one must observe that the axiom of reducibility appears in different mathematical systems under different names and in different forms, for example in Zermelo’s system of set theory as the axiom of separation, in Hilbert’s systems in the form of recursion axioms, and so on.

26 February to June 1939 Gödel co-taught two logic courses with Karl Menger at Notre Dame, one on set theory and the other more introductory. His introductory exposition was drawn from Russell, emphasizing the calculus of classes, the logical paradoxes, and Russell’s theory of types. At the same time, Gödel stressed that Russell’s axiomatization may and should be made “as simple as possible”; Menger recalls Gödel’s preoccupations with Leibniz as having been in evidence at this time (Dawson 1997: p. 136). Reference is made to the Notre Dame lectures at *Max Phil IX* [48] (cf. §6.3).

This late engagement with Russell's mathematical logic of a quarter of a century earlier is all the more remarkable because of the central place accorded the Axiom of Reducibility. While the axiom was considered redundant in their views of *Principia* by Wittgenstein, Ramsey, and later Quine, Gödel resurrected it as a key motif once one proceeds into the transfinite. Actually, it is the extent of the ordinals as given by the Axiom of Replacement that yields Gödel's "fundamental theorem", and so axiomatically, it is not Zermelo's Separation but Replacement, especially in its later guise as a reflection principle, that carries forth Gödel's affirmation of Reducibility.

In his lecture (1940a), Gödel started from the vantage point of trying to establish the consistency of the Continuum Hypothesis and again presented the  $L_\alpha$  hierarchy as "the ramified theory of types if extended to transfinite orders", pointing out how an "axiom of reducibility for transfinite orders" can be established. This time however he brought in the recursive scheme of functionals from Hilbert's attempt (1926) to prove the Continuum Hypothesis. With this, Gödel gave a model of the *Principia* system, constructed as his (1931) system  $P$ , in which the Continuum Hypothesis holds. In this conscious recasting of type theory, the types were essentially coded versions of  $L_{\omega_{n+1}} - L_{\omega_n}$ . One sees in all this a remarkable thematic coherence, and recalls the (1931) footnote 48a.

Gödel's next publication after those on  $L$  was his essay "Russell's mathematical logic" (1944). This would be Gödel's first philosophical publication, and his only extended publication on someone else's work—a measure, once again, of his engagement with Russell. It would also be Gödel's first robust statement of his platonism, and this arguably makes it a strong counterpoint in reaction against Russell.

Complaints about the inaccuracy of Gödel's reading of Russell have been lodged<sup>27</sup> but, as Bernays wrote (1946, p. 75), Gödel's main criticism of Russell's procedure in *Principia* is that it fails *clearly* to distinguish between the requirements for a general solution of the paradoxes and one tailored to a constructive or logically reconstructive point of view. Moreover, as described in §§3-5 below, there are more things going on in Russell's philosophy than a constructivist reading of *Principia's* theory of types can dream of, and of this Gödel was certainly aware.

Reviewing Russell's early work, Gödel (1944) got to the vicious-circle principle "no totality can contain members definable only in terms of this totality or members involving or presupposing this totality" (p. 133). He soon pointed out that this trifurcates into three principles corresponding to "definable only in terms of", "involving" and "presupposing" (p. 135). Only the first precludes impredicatively defined objects and the first "only applies if the entities involved are constructed by ourselves" (p. 136). He continued: "If, however, it is a question of objects that exist independently of our constructions, there is nothing in the least absurd in the existence of totalities containing members which can be described (i.e., uniquely characterized) only by reference to the totality". And this contradicts neither the second nor third forms, appropriately construed, of the vicious circle principle. Thus, Gödel significantly refined the issue of impredicativity, which in substantial part had led to his (1933o: p. 19) caveat about platonism mentioned earlier.

<sup>27</sup> Goldfarb (1989), Landini (1998: p. 8, 276), Parsons (1990: p. 110).

At the same time, there is a revisiting of Russell's view of the nature of truth as "the fundamental question" of philosophy. Gödel newly declared his realism extending to "concepts", their indispensability in mathematics to be regarded as a crucial factor (1944: p. 137):

Classes and concepts may, however, also be conceived as real objects, namely classes as "pluralities of things" or as structures consisting of a plurality of things and concepts as the properties and relations of things existing independently of our definitions and constructions.

What is the basis of this conception, according to Gödel? He appealed to a Russellian notion of "data", using scare quotes to mark the allusion (p. 137):

It seems to me that the assumption of such objects is quite as legitimate as the assumption of physical bodies and there is quite as much reason to believe in their existence. They are in the same sense necessary to obtain a satisfactory system of mathematics as physical bodies are necessary for a satisfactory theory of our sense perceptions and in both cases it is impossible to interpret the propositions one wants to assert about these entities as propositions about the 'data', i.e., in the latter case the actually occurring sense perceptions.

We shall return (§3) to the significance of the phrase "actually occurring sense perceptions", for it was chosen by Gödel with care. His remark resonates with Russell's idea of replacing *inference* (to independently existing bodies from actual sense-perceptions) with logico-mathematical *constructions* out of known particulars (in, e.g., *Our Knowledge of the External World* (1914b)). But it raises a more general question about truth. Russell had long used the notion of "data" to refer to the starting points, the residues exposed by philosophical analysis, what remains undefined at the end. So for Russell, "data" referred not only, and not primarily, to sense-perception *per se*, but also to primitive particular concepts.

The "data" are thus what is accepted as ultimate insofar as clarification has progressed, and it is clear that they are shaped by strictly formal developments, though not exhausted by them. After all, the indefinables themselves are known, relative to any given analysis, by non-demonstrative means.<sup>28</sup> On the very first page of the Preface to *Principia* Russell had adverted to this notion of "data" when he introduced the notion of analysis, regressive and progressive, in the context of the method of construction of a deductive system. The analogy with perception was then developed in *Principia* in the very presentation of *n*th truth we have already broached (cf. §§4.1-4.2).

It becomes clear enough from (1944) that for Gödel concepts can be taken to be Russellian propositional functions construed independently of their combinatorial, linguistic renditions and that for these there would be intensional paradoxes. Gödel did acknowledge (p. 139) that "there will exist properties  $\varphi$  such that  $\varphi(a)$  consists in a certain state of affairs involving all properties (including  $\varphi$  itself and properties defined in terms of  $\varphi$ ), which would mean that the vicious circle principle does not hold even in its second form for concepts". However, he pointed out that "this only makes it impossible to construct their meaning (i.e., explain it as an assertion about sense perceptions or any other non-conceptual entities), which is no objection for

28 Cf. Russell (1903): §§19,83), §83, stating that the notion of a *propositional function* has been revealed as an "ultimate" *datum*. Cf. also Russell (1908).

one who takes the realistic standpoint”.<sup>29</sup> He next noted (p. 140) that there would have to be “another solution to the paradoxes, according to which the fallacy (i.e., the erroneous axiom) does not consist in the assumption of certain self-reflexivities of the primitive terms but in other assumptions about these”. On the other hand, for “classes in the sense of pluralities or totalities”, that the vicious circle principle applies in the second form is “a very plausible assumption”.

Elaborating on this, Gödel turned to Russell’s constructivistic attempt to do without classes and concepts as real entities in the *Principia*, but observed obscurities, noting (p. 142) that “(primitive predicates and relations such as ‘red’ or ‘colder’) must apparently be considered as real objects”, and thus “neither the basic domain of propositions in terms of which finally everything is to be interpreted, nor the method of interpretation is as clear as in the case of classes”. These remarks resonate with Gödel’s *Max Phil* explorations in connection with Russell’s base case of the “simplest kind of *truth* and *falsehood*” (cf. §4 below). The underlying question is whether the notion of a propositional function might be analyzed in terms of a general intensional theory of concepts that could bear the weight of Gödel’s sophisticated elaboration of the notion of definability.

Gödel did assert (p. 143) that “What one can obtain on the basics of the constructivistic attitude is the theory of orders”. He noted that in the first edition of the *Principia* “the constructivistic attitude was, for the most part, abandoned” because of the reliance on the Axiom of Reducibility. He went on (p. 143f):

In the second edition of *Principia* (or, to be more exact, in the introduction to it) the constructivistic attitude is resumed again. The axiom of reducibility is dropped, and it is stated explicitly that all primitive predicates belong to the lowest type and that the only purpose of variables (and evidently also of constants) of higher orders and types is to make it possible to assert more complicated truth-functions of atomic propositions, which is only another way of saying that the higher types and orders are solely a *façon de parler*. This statement at the same time informs us what kind of propositions the basis of the theory is to consist, namely truth-functions of atomic propositions.

Note that, again, Gödel gets to the Russellian base case, here, the “atomic propositions”, this time through a constructivistic interpretation of Russell’s orders. We shall say more about how this is reflected in *Max Phil* IX and in Russell in §§3-5.

Gödel proceeded to voice a difficulty with this approach unless the number of individuals and primitive predicates is finite, and even if one confines oneself as did Russell “to such truth-functions as can actually be constructed”, there remains for the build-up of mathematics the (Poincaré-type) objection that propositional functions of order  $n$  defined syntactically presuppose arithmetic.

There is a thematic arc here going all the way back to first truth, second truth, etc. of the first quoted *Principia* passage of §1. Truth would be implicated in the definition of propositional functions of order  $n$  if they are to be taken as expressed in terms of predicates, linguistic expressions with quantifiers of order  $< n$ . As we had noted, first truth, second truth, etc. were woven into language in an informally recursive way

<sup>29</sup> *Max Phil* IX [48] makes a point that reinforces this one: “the concept of ‘definable concept’ is not defined” (full remark quoted below, §4.4). The translations of the *Max Phil* remarks in this article are by Juliet Floyd.

of thinking. However, a formalization of this would presuppose arithmetic—in fact Gödel's arithmetization of syntax!

Gödel continued, urging (at least as a first step) more incisive attention to the mathematical, as opposed to the more purely philosophical or dogmatic approach (p. 146f):

The theory of orders proves more fruitful if considered from a purely mathematical standpoint, independently of the philosophical question whether impredicative definitions are admissible. Viewed in this manner, i.e., as a theory built up within the framework of ordinary mathematics, where impredicative definitions are admitted, there is no objection to extending it to arbitrarily high transfinite orders. Even if one rejects impredicative definitions, there would, I think, be no objection to extend it to such transfinite ordinals as can be constructed within the framework of finite orders. The theory in itself seems to demand such an extension since it leads automatically to the consideration of functions in whose definition one refers to all functions of finite orders, and these would be functions of order  $\omega$ . Admitting transfinite orders, an axiom of reducibility can be proved. This, however, offers no help to the original purpose of the theory, because the ordinal  $\alpha$ —such that every propositional function is extensionally equivalent to a function of order  $\alpha$ —is so great, that it presupposes impredicative totalities. Nevertheless, so much can be accomplished in this way, that all impredicativities are reduced to one special kind, namely the existence of certain large ordinal numbers (or well-ordered sets) and the validity of recursive reasoning for them. In particular, the existence of a well-ordered set, of order type  $\omega_1$  already suffices for the theory of real numbers. In addition this transfinite theorem of reducibility permits the proof of the consistency of the axiom of choice, of Cantor's continuum hypothesis and even of the generalized continuum hypothesis (which says that there exists no cardinal number between the power of any arbitrary set and the power of the set of its subsets) with the axioms of set theory as well as of *Principia*.

This is a remarkably synthetic statement which happens to frame Gödel's mathematical progress in definability in terms of Russellian moves and incentives. Gödel initially shifted the ground to the possibilities of a *mathematical* investigation, where extensions carry the day, and the implicated impredicative definitions as the gauge of possible extent for the theory of orders. “Even if one rejects impredicative definitions, there would, I think, be no objection to extend it to such transfinite ordinals as can be constructed within the framework of finite orders”. One thus extends to  $\omega$  (shades of (1931) footnote 48a) and succeeding ordinals in autonomous progression (shades of (1933o)). In the 1960s, well-known mathematical work would investigate and delimit autonomous progression for “predicative analysis”<sup>30</sup>, and Gödel himself would probe autonomous progressions toward setting the limits to “finitary reasoning”.<sup>31</sup>

Gödel next elliptically alluded to his work on *L*, beginning with how in particular an axiom of reducibility can be proved with transfinite orders although this requires an ordinal so great that it “presupposes impredicative totalities”. “Nevertheless, so much can be accomplished in this way, that all impredicativities can be reduced to one special kind, namely the existence of certain large ordinal numbers (or well-ordered sets) and the validity of recursive reasoning for them”. With this dramatic view of *L* as providing a reduction of *all* impredicativities to the extent of the ordinals, Gödel went

30 Cf. Feferman (1962), Feferman (1964), Schütte (1965a), Schütte (1965b).

31 Cf. van Atten (2014).

on to elliptically describe the workings of “the transfinite theorem of reducibility” for getting relative consistency.

In the large, this passage coheres with our nuanced reading of Gödel’s mathematical platonism as emergent and bears on its operative significance. In the lecture (1933o) Gödel had broached autonomous progression while expressing dissatisfaction with how the axioms allow impredicative definitions. Now the impredicative axioms of set theory are to be realized, and this is made possible because an axiom of reducibility is at work. Gödel, having made the jump sometime in 1933–1935 to entertaining arbitrary ordinals as given by Replacement, duly took on the inherent axiomatic impredicativity. Gödel’s newly declared realism in (1944) extends moreover to concepts, and this is an espousal which can in part be seen through the essay as his commitment to truth, especially for the primitive or simple concepts which form the basis of the Russellian treatment.

As for operative significance, Gödel’s mathematical platonism would not be a naïve affirmation but would stabilize specifically on the need for metamathematics, broadly construed, and direct transfinite reasoning, e.g. truth for arithmetic and the extent of the ordinals for  $L$ , and how for this the reality of the ordinals and the like are indispensable.<sup>32</sup> Be this as it may, once the stage has been set through realist advocacy, Gödel would nonetheless be perennially interested in seeing how constructive hypotheses and approaches would go. For example, with an evident adumbration in the concluding remarks of (1933o) involving intuitionistic mathematics, Gödel in 1941 developed proof-theoretic ideas for establishing the consistency of number theory in terms of a logic-free theory of functionals of finite type, this eventually being published (1958), the “*Dialectica* interpretation”.<sup>33</sup>

With his mathematical platonism declared and stabilized, Gödel in the next succeeding years would expand on truth and definability in set theory.<sup>34</sup> In oral, brief but pithy remarks, Gödel (1946) explored possible absolute notions of demonstrability and definability, these not to be dependent on any particular formalism. For absolute demonstrability, Gödel entertained the heuristic of *reflection*, in evident recall of the Axiom of Reducibility, and *large cardinal axioms*, suggesting that the following could be true: “Any proof for a set-theoretic theorem in the next higher system above set theory (i.e. any proof involving the concept of truth which I just used) is replaceable by a proof from such an axiom of infinity”. For absolute definability, Gödel entertained

32 In that letter of 7 December 1967 to Wang, after pointing out how Skolem did not get to the Completeness Theorem as it required non-finitary reasoning, Gödel wrote (2003b: p. 396ff):

... I may add that my objectivist conception of mathematics and metamathematics in general, and of transfinite reasoning in particular, was fundamental also to my other work in logic.

How indeed could one think of *expressing* metamathematics *in* the mathematical systems themselves, if the latter are considered to consist of meaningless symbols which acquire some substitute meaning only *through* metamathematics?

Or how could one give a consistency proof for the continuum hypothesis by means of my transfinite model  $\Delta[L]$  if consistency proofs have to be finitary?

33 As late as 1974, in a conversation with Sue Toledo (2011), we find Gödel discussing in a remarkably open-minded way the prospects for finitary mathematics; the tenor of Gödel’s platonism, as Curtis Franks (2011) has noted, is self-critical and knowledge-oriented, rather than dogmatic.

34 For elaboration on what follows, see Kanamori (2007).

a transfinite hierarchy of “concepts of definability”, expanding language to allow constants for every ordinal, and pointed out that “by introducing the notion of truth for this whole transfinite language, i.e., by going over to the next language, you will obtain no new definable sets (although you will obtain new definable properties of sets)”. Thus is the unity and content of truth unfolded in its progress through the hierarchy.

In his well-known article on the continuum problem (1947), he opined that the Continuum Hypothesis would be undecidable as per the axioms of set theory, and entertained the possibility of large cardinal axioms “based on hitherto unknown principles” that might decide the Continuum Hypothesis and other set-theoretic questions. In an oft-quoted passage, Gödel argued that for a new axiom “a decision about its truth is possible” merely based on its “success” according to “fruitfulness in consequences”, and that there might exist axioms moreover “shedding so much light on a whole discipline, and furnishing such powerful methods for solving given problems . . . that . . . they would have to be assumed” —an imagined indispensability argument for reaching beyond the axioms of set theory and setting the “data” into an ultimately correct frame.

Gödel concluded (1947) by forwarding the remarkable opinion that the Continuum Hypothesis “will turn out to be wrong” since it has as paradoxical consequences the existence of thin, in various senses articulated, sets of reals of the power of the continuum. These examples evidenced how Gödel was aware of the work of the descriptive set theorists in articulating the nature of the continuum. On this, it is especially notable that decades later Gödel in handwritten notes (1970a, 1970b) proposed “orders of growth” axioms, related to classic work of Felix Hausdorff, for deciding the value of the continuum. However, “Some considerations leading to the probable conclusion that the true power of the continuum is  $\aleph_2$ ” (1970a), and “A proof of Cantor’s continuum hypothesis from a highly plausible axiom about orders of growth” (1970b), while in the spirit of taking on new axioms to decide the “true” power of the continuum, drew contrary conclusions! In (1947) Gödel may have cast a light on the Continuum Hypothesis from a platonistic platform, but as before, once the stage has been set he was willing to go wherever mathematical analysis would take him.

Returning to (1944), Gödel concluded his essay with a short discussion of the simple theory of types and the concept of analyticity, especially in connection with concepts. After pointing out the importance of the ambiguity of the possible arguments for proposition functions, Gödel wrote (p. 148):

If on the other hand one considers concepts as real objects, the theory of simple types is not very plausible since what one would expect to be a concept (such as, e.g., ‘transitivity’ or the number two) would seem to be something behind all its various ‘realizations’ on the different levels and therefore does not exist according to the theory of types.

Gödel then speculated (p. 150) about prospects for a theory of concepts and possible solutions to intensional paradoxes. Distinguishing the notions of “tautological” and “analytic” from one another (p. 151, n. 46), he pointed out that to hold that the truths of logic and mathematics are conceptual, i.e., reducible to truths about identities of concepts, does not eliminate the idea that their axioms have content “if the reduction

is effected not in virtue of the definitions of the terms occurring, but in virtue of their meaning, which can never be completely expressed in a set of formal rules” (p. 151, n. 47). Thus he devolved back to the issue of what remains indefinable after analysis and axiomatization, invoking a kind of “unsaturatedness” with respect to the primitives.

Recalling that in the second edition of *Principia* “Russell took the course of considering both classes and concepts (except the logically uninteresting primitive predicates) as nonexistent and of replacing them by constructions of our own” (p. 152), Gödel noted the limited strength of the results and commended reconstructive efforts (e.g., the suitably reinterpreted *first* edition of *Principia*, using the simple theory of types) taking “a more conservative course” of “trying to make the meaning of the terms ‘class’ and ‘concept’ clearer, and to set up a consistent theory of classes and concepts as objectively existing entities”. Beginning his transition to the closing, speculative paragraphs of his essay, Gödel then remarked that “many symptoms show only too clearly, however, that the primitive concepts need further elucidation”.

Gödel’s closing remarks in (1944) offered the optimistic idea that a newly fashioned, realistic concept theory might be what is needed to “elucidate” the foundations. This, he suggested, could realize Leibniz’s dream of a universal characteristic (p. 152):

It seems reasonable to suspect that it is this incomplete understanding of the foundations which is responsible for the fact that mathematical logic has up to now remained so far behind the high expectations of Peano and others who (in accordance with Leibniz’s claims) had hoped that it would facilitate theoretical mathematics to the same extent as the decimal system of numbers has facilitated numerical computations. For how can one expect to solve mathematical problems systematically by mere analysis of the concepts occurring if our analysis so far does not even suffice to set up the axioms? But there is no need to give up hope.

Leibniz himself, Gödel reported (p. 153), expressed the hope that a few “select scientists” might be able to develop logic “to such an extent” (Gödel next quotes from Leibniz directly) “that humanity would have a new kind of instrument increasing the power of reason far more than any optical instrument has ever aided the power of vision”. Gödel noted, correctly, that in this same passage Leibniz had estimated that the work would take about five years to complete.

As Parsons has written (1990: p. 118), these purely speculative remarks, forming a kind of *coda* to (1944), are perhaps the “most striking and enigmatic utterances” in all of Gödel’s published writings. They may seem to be tacked on at the end of the essay, poorly motivated, and they are certainly not defended, offering instead a profession of the rational optimism for which Gödel is still known.<sup>35</sup> More generally, throughout (1944) Gödel did not follow Russell in detail, but offered, as Parsons observed (1990: p. 110, n. *m*), a “conscious assimilation of Russell’s conceptual scheme to his own”, rather than a straightforward exposition of Russell’s philosophy.

However, the *Max Phil* notebooks IX–X Gödel kept while writing (1944) shed new light on Gödel’s *coda*. From them, one learns that the scope of Gödel’s sense of philosophy was, in fact, as broad as Russell’s, and deeply indebted to Russell—something one cannot know on the basis of Gödel’s publications and published correspondence alone. Not only the disagreements, but the intellectual

<sup>35</sup> Cf. van Atten and Kennedy (2003). One might add that it is part of mathematical culture, as it is not of philosophical culture, to finish with a description of open problems and future work.

kinship and inspiration with Russell should be acknowledged. In Russell Gödel felt he saw a philosopher with whom he could find agreement on certain essentials, especially with regard to the concept of truth. In particular, when set beside Russell's purely philosophical writings about truth, his remarks look far more *à propos*—if also, it must be said, more philosophically ambitious.

What did Gödel know or believe about Russell's philosophical evolution? Our next section (§3) will pursue this question with regard to Russell's purely philosophical writings. As for Russell's work in mathematical logic, we know that among Russell's pre-*Principia* writings there is no question that Gödel knew (1906a) and (1906b), where Russell had proposed a unitary approach to the paradoxes of the transfinite and broached his “no-class” theory. Both are explicitly referred to by Gödel (1944: p. 131, n. 9; p. 141, n. 31), and he even pointed out (p. 131, n. 9) that (1906b) was “Russell's first paper” on the issue of whether “every propositional function exists ‘as a separate entity’” —alluding to Russell's own phraseology of “entities” in (1906a) and (1906b).

Yet immediately thereafter, continuing his thought, Gödel remarked (p. 131, n. 9): “If one wants to bring such paradoxes as ‘the liar’ under his [Russell's] viewpoint, universal (and existential) propositions must be considered to involve the class of objects to which they refer”. Russell (1906b) did not, in fact, contain any explicit remark about the Liar, as it focused on paradoxes of the transfinite, including difficulties about the Axiom of Choice. The Liar, unlike these paradoxes, entangles us with the concept of truth. So Gödel's note was intended to make a further point: in order for Russell to carry off his earliest approach to the paradoxes, all the way back down to the foundations or “commencements de la Logistique”, as Russell called them (1906a: p. 649), he would have to adopt a form of realism about propositional functions or concepts. As we have seen, Gödel had noted (1944: p. 141) that not *every* move could be interpreted as a mere “*façon de parler*”.

This coheres with the next step Gödel made in (1944), which was to advert to Russell's discussion of “judgments of perception” in *Principia*, the *Ur*-point at which “the simplest kind of *truth* and *falsehood*” had been defined (*Principia*: p. 43). As to ultimate beginnings and endings, Gödel was correct that “the method of interpretation” was hardly clear. For in *Principia*, at just the point where “finally everything is to be interpreted”, Russell adverted to his correspondence theory of truth, the “multiple relation theory of judgment” (MRTJ).

As we shall argue in what follows, Gödel's remark about primitive perceptual predicates in *Principia* was not placed in (1944) simply to note Russell's exercise in similarity-relation constructions from sense data with the method of the ancestral.<sup>36</sup> More fundamentally, Gödel was turning toward Russell's own treatment of the nature of truth, the MRTJ. In the remainder of this paper we shall unfold some of the purely philosophical features of this backdrop, for these shaped the framework within which Gödel wrote *vis-à-vis* Russell.

36 Although Gödel (1944: p. 142, n. 32) is careful to refer to Appendix C of *Principia*.

### 3. Philosophy and truth

How precisely *did* Gödel make the transition from mathematical logic to philosophy? This question cannot be answered by situating him against the broad backdrop of the history of philosophy alone—although recent scholarship has established that Gödel did have a better understanding of this than one might have supposed,<sup>37</sup> and *Max Phil* IX-X contain, in addition to discussion of mathematical logicians, a wide range of remarks about Plato, Aquinas, Leibniz, theology, psychology, and physics. Nevertheless there remains an unfortunate tendency to suppose that Gödel leapt, in a kind of ungrounded formal analogy, from incompleteness to certain general doctrines, which he then rationalized by appeal to philosophers he had read.

This is not so. An obvious point is that he was quite aware that Carnap had drawn philosophical conclusions from the incompleteness proof that differed very much from his own. Even more, however—if one attends to Gödel’s own mathematical practice and understands how important was his immersion in *Principia* and Russell’s philosophical writings in the early 1930’s—one can see at work in the diffuse and complicated workings of his *Max Phil* notebooks IX-X a relatively informed and sophisticated sequence of ideas about how philosophy and foundations of mathematics might be seen to work in a dialectical interplay with one another.

Many questions remain about how Gödel conceived his philosophical writings. We shall argue that Russell was a crucial source and inspiration, at least in 1942–1943. The arguments in support of our interpretation will be multi-faceted. In this section we provide a thematic overview and guide to the remainder of our paper, in which we emphasize subtleties involved in Gödel’s way of working: his knowledge of Russell’s history, the space of philosophical possibility he saw before him, and his tenacious, imaginative, and persistent drive to rummage anywhere and everywhere he could to find fruitful, yet fundamental ways of thinking.

The upshot of our reading will be to suggest that at the end of (1944), in proposing the development of a theory of concepts in the tradition of Leibniz, Gödel had in mind an *infinitary* form of Russell’s MRTJ: that is, a correspondence theory of truth made out in terms of reflective containment (or projectability) of one (infinitely complex) conceptual complex inside another. This fits, at least in a kind of vivid pictorial way, Gödel’s practice in providing relative consistency proofs in set theory *via* reflection. More interestingly, as we shall explain (§4), his knowledge of the MRTJ would have derived from precisely the same *Principia* passage about *n*th truth we have already discussed, the one that got him to the undefinability of truth on the way to incompleteness in 1930 (cf. §1).

At the end of (1944), although Gödel neither spoke explicitly of truth as correspondence<sup>38</sup> nor named Wittgenstein explicitly, he did suggest an infinitary theory of concepts as the culmination of a demand that the theory of descriptions not be seen as reducing all singular reference to a mere “*façon de parler*”. There is little doubt that Gödel felt Wittgenstein had influenced Russell too far in the direction of construc-

<sup>37</sup> Wang (1987, 1996), Parsons (1995, 2010), van Atten and Kennedy (2003), Kennedy and van Atten (2009), Toledo (2011), Crocco (2006, 2012), Franks (2011).

<sup>38</sup> In general Gödel was disinclined to use the notion of truth, wishing to avoid controversy.

tivism. The infinitary version of the MRTJ was intended by Gödel, we suggest, to turn Russell away from that influence and in a very different philosophical direction.

Gödel's philosophical picture of truth for mathematics remained, philosophically speaking, a glittering vision in the distance. It does not appear in (1944), because Gödel did not manage to work out a theory of truth, and he knew this. What is important, however—given the understandable skepticism about Gödel's philosophy expressed in much recent literature—is that he understood that to fashion and defend a substantial correspondence theory of truth for the infinite would require systemic reflection on the notions of conceptual complex, perception, meaning and truth, a careful reworking of the analogies among them, and in a sufficiently convincing way. His philosophical ambitions in making this move were enormous, and he knew what the kinds of constraints on argumentation would be, if he followed Russell's path from the first edition of *Principia*, Introduction, chapter II.

Hints of Gödel's ambitions remained in (1944), such as the point (p. 130) where he remarked explicitly, of the problem of “collapsing arguments” —what Russell called in (1940) the “indication” of a sentence, in response to the difficulties he had always had with Frege's view of sentences as names of truth-values—that “I cannot help feeling that the problem raised by Frege's puzzling conclusion has only been evaded by Russell's theory of descriptions and that there is something behind it which is not yet completely understood”.<sup>39</sup> The infinitary form of the MRTJ, with infinitely complex conceptual complexes, was a stab in the direction of a radically new way of approaching meaning and truth.

A stab is not a successful theory. Our point, however, is that Gödel had a better sense of how purely philosophical handling of the notions of truth and meaning would have to work than can be gleaned from his published writings. He knew that it was an undeveloped stab. Yet, following Russell, he conceived his move outside mathematics to pure philosophy, and the theory of perception, as absolutely necessary in light of the notion of truth. Conversely, it was the notion of truth that his analogy with sense-perception was intended to illuminate, and not simply (for example) the notion of *set*.

Our analysis will show that the road to Gödel's coda to (1944) was paved by thorough study of Russell's writings, both mathematical *and* philosophical. Given his temperament and careful ways of working, it is no surprise that Gödel was a systematic and not merely a narrow reader of Russell's writings.<sup>40</sup> This, judging from *Max Phil* IX-X, shaped his (1944) response to Russell. He was aware, for example, that for years after writing *Principia*, Russell sought to fashion a theory of knowledge to fit his logic. Just here is where Gödel believed Russell had, under the influence of Wittgenstein, gone wrong.

Of course, Gödel was in 1942-1943 unaware of Wittgenstein's evolution away from the *Tractatus*, and the uncanny fact that he too, having pondered Turing's analysis of the general notion of formal system and read Russell's later writings, was also writing a series of notebooks (1937-1944) revisiting the philosophical

39 Cf. Neale (1995) for a reconstruction of Gödel's “collapsing” argument that turns in the direction of philosophical logic and linguistics.

40 We have noted that Gödel (1944: p. 129, n. 7,137) refer to Russell (1940). In reprinting (1944) Gödel added the full title of Russell (1940) to the first occurrence to make his reference crystal clear (cf. [1990: p. 315]).

foundations of *Principia*. It is no surprise, when we ponder this fact, that we see Gödel revisiting, in *Max Phil IX-X*, some of the very same issues known to readers of Wittgenstein’s *Remarks on the Foundations of Mathematics*: the “feel” of a language; the contrast between following a rule and taking in a sense of the meaning of what one is doing<sup>41</sup>; the nature of aspect-perception and consciousness; the problem of confusing truth with correctness; the surveyability of proofs.<sup>42</sup> Both of these thinkers’ writings were drawn from the same sources.<sup>43</sup> Moreover, neither took Tarski to have resolved philosophical questions about truth. In this they each followed Russell.

Let us take stock of our characterization of Gödel’s mathematical practice. So far we have seen Gödel steadily exploiting the interplay between truth and definability—using undefinability and definability at various levels, along with the heuristic of reflection—as a way of proceeding that unfolds truth by delivering relative consistency proofs along the way. The pattern Gödel espoused is this: define where possible and axiomatize; transcend; look back to reflect on and analyze the transcendence; define and axiomatize where possible again, reflecting on means and connections; and continue transcending as far as possible. If truth is attained at all, it is attained in the course of this progress through the absoluteness of notions such as “formal system” and “definable”.<sup>44</sup>

Gödel emphasized throughout (1944) the ways in which his use of a transfinite Axiom of Reducibility would demand that the theory of descriptions not amount, in the end, to a merely linguistic device. But, after all, definability was part and parcel of his mathematical practice with relative consistency proofs in *L*, providing a kind of “inner” correspondence and fertile use of descriptions. From Gödel’s point of view,

41 We read at *Max Phil IX* [16]:

A board game is something purely formal, but in order to play well, one must grasp the corresponding content [the opposite of combinatorically]. On the other hand the formalization is necessary for control [note: for only it is objectively exact], therefore knowledge is an interplay between form and content.

42 The opening of *Max Phil IX* is striking in this respect:

*Remark (Grammar)*: The calculus in which one speaks should be simple and therefore contain many *definitions* [especially the  $\exists$ ]. The calculus about which one speaks need not be simple at all, but “easy to describe”, therefore containing possibly very few fundamental signs and defined signs.

43 Wittgenstein’s 1937 writings on Gödel’s proof of incompleteness (1978: I App. III [1937]) are uneven and were never intended for publication. It is unsurprising that Gödel wrote to Bernays in 1958 (2003a: p. 161) that it seemed to him that “the benefit created” by their publication “may be mainly that it shows the falsity of the assertions set forth in it. [Note: And in the *Tractatus*. (The book itself really contains very few assertions.)]”. When Gödel wrote to Menger in 1972 to comment on Wittgenstein’s remarks on his theorem in the 1956 edition, he was also very critical, stating (2003b: p. 132-3) that the remarks were extremely weak, that Wittgenstein either didn’t understand incompleteness or “pretended” not to (cf. Wang 1987: p. 49). Interestingly, Gödel made clear he wished to contradict what he took to be Wittgenstein’s main point: that incompleteness was a *mere* paradox—i.e., one not yielding anything of fundamental interest about the notion of truth. However, Wittgenstein’s understanding appears to have progressed somewhat over time, especially by 1940, after discussions with Turing and Kreisel (cf. 1978 VII, §§19ff. [1941]), Floyd (1995, 2001) and Floyd and Putnam (2012)).

44 Kennedy (2013) develops a notion of “formalism freeness” to show how Gödel’s standpoint continues to bear fruit mathematically; cf. Kennedy (2014), Kennedy (2015) on the idea of “border crossing” among a variety of logical frameworks as a philosophically and mathematically rich perspective.

the way to avoid a collapse of the whole interpretation of the theory was to propose an infinitary “concept theory” constituting a correspondence theory of truth suiting relative consistency proofs that take the transfinite ordinals as given.

Throughout his career Gödel remained ecumenical, deeply interested in comparing proof procedures and assumptions, even constructivist ones, to see what they could do. For example, he never offered an across-the-board defense of the law of the excluded middle, but instead explored the nature and limits of embracing it and delimiting its sphere of applicability. Ultimately, the “law” of the excluded middle comes down to this: either one accepts it, or one does not. A formal rule may be added to a system, or not. Such is the *neutral* character of what is scientific, to borrow a term from Russell we shall encounter again (§5.1). Yet in making what Gödel called—in a Russellian vein—an “ethical legitimation” of such a choice (*Max Phil* IX [41]), one requires philosophy.<sup>45</sup>

Thus transcendence was not all for Gödel. Neither truth (ontology) nor proof (epistemology) are, in general, definable or reducible one to the other, nor are these notions to be analyzed without residue by purely logico-mathematical means. Their interplay is therefore at least partly a purely philosophical matter, for Gödel. As we see in broaching more purely philosophical aspects of his work, Gödel embraced a self-imposed requirement to continually revisit his primitive notions and assumptions afresh, to draw out new mathematical aspects and analyses that enriched the perceived basis at the roots.

But of course a key question is this: In what sense *is* the basis “perceived”? Not, of course, *literally* in sense-perception, as Gödel was, at least in 1942-1943, fully aware. Quite understandably, there has been much skeptical reaction to Gödel’s remarks about sense-perception and concept realism in his published writings.<sup>46</sup> How is such an analogy to be worked out? Just here it is worthwhile to emphasize that Russell provided Gödel with specific inspiration. Obviously, if Gödel was to take the infinite as in some way “given”, the issue at hand was, not so much *that* it is a given, but *how*. And just here one detects the impact of Russell on Gödel’s philosophy.

The analogy with sense-perception figured quite explicitly in getting Russell *onto* the hierarchy of truth in *Principia*, so that he could *begin* with “first truth”. This required him to define “the simplest kind of *truth* and *falsehood*” (p. 43), as we have seen. At just this place in his presentation of ramified type theory—facing “*the* fundamental question [of philosophy], namely, the nature of truth”<sup>47</sup>—Russell resorted to pure philosophy, devolving back upon what he called “judgments of perception”. This Gödel saw, and probably understood, even back in his salad days in Vienna.

But what, then, did Russell *mean* in devolving back upon “judgments of perception”? Was this not empiricism, a form of sense-data theory, run amok? Not necessarily—or anyway, not obviously—as Gödel himself helps us to see. As we shall soon explain, in *Principia* Russell adverted to “judgments of perception” for philosophical and logical

45 Russell’s (1940: ch. XX) is on the law of the excluded middle, and as Gödel would have noticed, Russell stated there (p. 288) that we “may accept” the law because we “need” it, having determined that “‘truth’ must be a wider conception than ‘knowledge’”. For a defense of Wittgenstein’s insistence that Gödel’s incompleteness theorem demands philosophy to be understood, see Floyd and Putnam (2012).

46 Parsons (2010), Burgess (2014).

47 Again, Russell (1908).

reasons, and not in the first instance for psychological or epistemological reasons—and the notion at issue here was that of *truth*.

Gödel opened (1944) with a lament at the lack of formal precision in the *Principia*, stressing that he did not wish to focus the bulk of his essay on either its formal or mathematical content.<sup>48</sup> Instead, he would focus on (p. 127):

... Russell's work concerning the analysis of the concepts and axioms underlying mathematical logic. In this field Russell had produced a great number of interesting ideas some of which are presented most clearly (or are contained only) in his earlier writings.

This reference to Russell's "earlier" writings must be taken very seriously, as we shall see. Gödel had a theory—and considerable knowledge—of Russell's evolution. In fact he justified his developmental picture of Russell by quoting (p. 127) his favorite Russell remark<sup>49</sup> (1919: p. 169):

Logic is concerned with the real world just as truly as zoology, though with its more abstract and general features.

The quotation was produced in (1944: p. 127) in order to praise Russell's "pronouncedly realistic attitude", although Gödel also remarked, more critically, that Russell's attitude "has been gradually decreasing in the course of time ... and ... always was stronger in theory than in practice".

In fact, the evidence shows that by 1932, and possibly even earlier, Gödel had acquainted himself, not only with Russell's "earlier" logical and mathematical work, but also with Russell's "earlier" purely philosophical writings (including *The Problems of Philosophy* (1912), *The Analysis of Mind* (1921), and *The Analysis of Matter* (1927a)), each including chapters concerned with the nature of truth and falsehood. Most importantly, he had a clear understanding of the transition from logic to philosophy that occurred in *Principia's* treatment of *n*th truth. And he had followed the development of these ideas in Russell's subsequent evolution.

A clue is that early in *Max Phil* IX [9-10], beginning a list of eight numbered observations, Gödel wrote:

1.) "Being" is a predicate of all things [note: to be understood as similar to the *Russellian* type signs [*Typenzeichen*]], just as "life" <is> of all is objects of zoology. Only we can't "look out" over Being as we can "look out" over life. 2.) The concept of a thing exists in a strong sense as the thing, even if this exists. But even non-existing, yet possible things have a "weak" existence as that which falls under the concept.

Gödel is here concerned to say that the treatment of logical notions such as "Being" (and "Truth") essentially involve notions of possible existence, as zoology does not. Thus any attempt to come to terms with these (in general) undefinable and unsurveyable notions will have to go beyond, not merely Russell's type theory as it stood in the second edition of *Principia*, but also his philosophy. "Being", or truth, requires a different method of conceptualization from that of "life", however fundamental the latter notion may be to zoology and our theory of nature.

<sup>48</sup> Cf. Wang (1987: p. 270).

<sup>49</sup> Cf. Wang (1987: p. 112).

Among Russell's philosophical essays that must have most interested Gödel in this regard were those representing the high watermark of Wittgenstein's impact on Russell, *The Philosophy of Logical Atomism* (1918). Russell gave these initially as a series of eight lectures in early 1918, before the prison stint during which he wrote *Introduction to Mathematical Philosophy* (1919a); they were subsequently published in a series of five articles in *The Monist* 1918-1919, which Gödel at some point read.<sup>50</sup> Their contents, Russell explained at the very beginning of (1918), were

... very largely concerned with explaining certain ideas which I learnt from my friend and former pupil Ludwig Wittgenstein. I have had no opportunity of knowing his views since August 1914, and I do not even know whether he is alive or dead. He has therefore no responsibility for what is said in these lectures beyond that of having originally supplied many of the theories contained in them.

Russell's analogy between logic and zoology was first laid out in Lecture IV of (1918), and precisely so as to defend "realism", just as Gödel would later say (1944: p. 127):

... I think one might describe philosophical logic, the philosophical portion of logic which is the portion that I am concerned with in these lectures since Christmas (1917) as an inventory, or if you like a more humble word, a "Zoo" containing all the different forms that facts may have. I should prefer to say "forms of facts" rather than "forms of propositions". . . In accordance with the sort of realistic bias that I should put into all study of metaphysics, I should always wish to be engaged in the investigation of some actual fact or set of facts, and it seems to me that that is so in logic just as much as it is in zoology. . . . Now I want to point out today that the facts that occur when one believes or wishes or wills have a different logical form from the atomic facts containing a single verb which I dealt with in my second lecture.

Gödel's knowledge of Lecture IV of (1918), as well as other of Russell's purely philosophical writings, helps to explain, not only the shape and tenor of (1944) itself, but also the—*prima facie* much more puzzling—discussions of consciousness, mind, action, emotion, will and belief that are strewn throughout *Max Phil IX-X*.

The gist of the argumentative situation, as Gödel perceived it, was this. Russell's conception of the theory of truth tended to be modeled, after 1910, by a finitary conception of the nature of the "multiple relation" of correspondence constituting the MRTJ, largely because of Russell's interest in what he called "the theory of symbolism" (1918: I, IV, VII). Russell was careful not to rule out the possibility of a more complicated version tied directly to metaphysics, as Gödel knew from reading Russell (1918: IV §4, VII) and (1940). Indeed, Russell discussed Gödel explicitly in connection with the issues of infinitary vs. finitary approaches in a chapter on "The Object-Language" in his William James Lectures (1940: p. 71), and, judging from *Max Phil IX-X*, Gödel had this book open before him while he was composing (1944).

In presenting the philosophical problem of the nature of truth and falsehood, Russell had, however, historically tended—and wrongly, from Gödel's point of view—to emphasize the insurmountable difficulties he had faced with the finitary version of the MRTJ. His failure in 1913 to find a satisfactory theory of knowledge to fit the quantifier-free basis for the MRTJ in "judgments of perception"—difficulties encountered during his period of strong engagement with Wittgenstein and his

50 See §6.1 below for the history of their publication and their availability to Gödel. Gödel quoted from IV §4 without citation at (1944: p. 129, n. 7).

abandonment of his (1913) manuscript on “The Theory of Knowledge”—led Russell to conclude, as he put it (1918: IV §3), that a “map-in-space” of belief cannot be drawn, a point he illustrated (more than once) with a finite geometric picture of relations among objects and minds.<sup>51</sup>

In (1918: IV) Russell attributed the “discovery of this fact” about a “map-in-space” to Wittgenstein. Colorfully putting the point about “judgments of perception” in terms of his zoological metaphor, Russell wrote, of judgments involving willing, wishing, and so on that (1918: IV§3): “I have got on here to a new sort of thing, a new beast for our Zoo, not another member of our former species but a new species”. Wittgenstein, of course, would have rejected this zoological gloss on his idea, for—as Gödel knew from reading the *Tractatus* during the heyday of its influence on the Schlick circle in Vienna—his point to Russell had been, not that there was a new form in view, but, rather, that a recasting of the whole notion of form as the *possibility* of structure would be needed, and this is just what Wittgenstein worked out in the *Tractatus* (1921: 2.033). Russell saw logic as a “skeletal” enterprise of classifying and identifying bones beneath beliefs, a framework “within which the test of coherence applies” (1912: ch. XII), whereas Wittgenstein took logic to be itself a “scaffolding” to be taken up and taken down, an aid in the construction of a true depiction of reality, but not itself in the business of studying what actually is (1921: 3.42, 4.023, 6.124). Russell’s effort to eliminate “possibility” in favor of actual correspondence failed in the face of Wittgenstein’s new conception of logic. And Russell granted this fact, ever after, not only for logic, but even for mathematics.

A Gödelian way to recast Russell’s remark about “our Zoo” would point elsewhere. Instead of rejecting the analogy with zoology altogether—as Wittgenstein had done—Gödel embraced it, while at the same time drawing from the *Tractatus* the idea of form as the possibility of structure. Gödel then drew a Russellian conclusion from the futility of Russell’s attempts to use a finitary formalization to analyze the nature of truth, belief and meaning, arguing that the notion of conceptual order must be taken as metaphysically and experientially basic—even if operationally controlled, insofar as possible, by the application of the axiomatic method in proof.

Thus in (1944) Gödel took himself to be endorsing just the kind of account based on “judgments of perception” in terms of the MRTJ that Russell had offered in (the first and second editions of) *Principia*—but this time incorporating into the ground theory infinitely complex conceptual complexes, containing within themselves a “reflective” and ordinal structure. If one takes the possibility of conceptual order—infinitary or otherwise—o be given from the start, Wittgenstein’s 1913 objections to Russell have no purchase, and internal, relative correspondence can be said to hold within the  $U_r$  level of Russellian  $n$ th truth. Moreover—most important of all for Gödel—a unified view of truth as correspondence could be retained, covering logical, mathematical and empirical truth.

In fact, Gödel’s transition to even wider philosophical ruminations on the nature of experience, perception, the self, emotion, consciousness, and meaning is quite comprehensible if we take Russell’s own philosophical evolution, as perceived by

51 For example, in (1918: IV), but also in (1913: ch. II, II) and its accompanying manuscripts. For a discussion of Russell’s diagrams, cf. Carey (2003); the wider context of Russell and Wittgenstein is covered in Carey (2007), cf. Zalabardo (2012).

Gödel, into account. In particular, Gödel always took it to be a natural step from *Principia's* treatment of *n*th truth to the foundation Russell provided there in terms of “judgments of perception”. Later on, in his William James Lectures (1940), Russell continued to attempt to ground the distinction between truth and falsehood in “judgments of perception”, rehearsing once again his arguments with Wittgenstein and developing an approach to truth that differed considerably from that of Tarski. Gödel's (1944) is in part a response to this.

Though Gödel (rightly) blamed Wittgenstein for causing Russell's retreat from the unified approach to truth as correspondence he pursued in *Principia*—as well as for the more constructivistic treatment of orders in its second edition—yet, having been so influenced by Russell's original treatment of “judgments of perception” and the MRTJ in *Principia*, as well as in Russell's subsequent publications, Gödel was, from 1932 onward, fascinated by Russell's long-lasting ambition to analyze the notions of *experience* and *belief*.

As a result of reading (Russell 1912, 1921, 1927a) in 1932, as well as Russell's (1940) in 1942-1943 (cf. §6.1), Gödel developed an interest in Russell's “neutral monism”, the doctrine, derived from William James, that consciousness does not exist as a separate quality, but is reducible to complexes of relations incorporating the subject's environment and body. In (1940) Russell revisited the notions of truth and meaning in light of neutral monism, and several of his arguments that metaphysics cannot be wholly avoided turned on the idea that the foundations of mathematics must be philosophically related to the foundations of psychology and the theory of meaning, but in such a way that the notion of truth is not reduced to a merely psychological or linguistic notion.

Delving into the unbounded complexity of experience and meaning, Gödel believed that his infinitary approach to concepts was just what Russell needed to regain a better foothold on the nature of mathematical truth. Ultimately, Gödel's perspective would lead to the transition, in the second half of *Max Phil X*, to his engagement with Leibniz's theory of monads—a return, in the end, to Russell's even “earlier” writings (Russell 1900)—in which the qualitative aspects of conscious experience are converted into an infinitary metaphysics doubling as a theory of truth. Our account stops at this point, having established the importance of Gödel's engagement with Russell in 1942-1943.

The structure of our story is as follows. We turn first to *Principia* and the MRTJ, analyzing the continuation of the passage we have already argued had such importance for Gödel's reaching the undefinability of truth in 1930 (§4.1). Next, we examine Russell's step into philosophy at the end of this passage, examining his idea of “judgments of perception” (§4.2). The analogy between grasp of simple notions and sense-perception is analyzed (§4.3), and then we review the history of Russell's struggle with *The Theory of Knowledge*, and Wittgenstein, in 1913 (§4.4). We attend throughout to Gödel's way of reading Russell.

We turn next to the later Russell of *An Inquiry into Meaning and Truth* (1940). We define and discuss the significance of “neutral monism” (§5.1) and then turn to Gödel's reworking of the view into an infinitary form of the MRTJ in *Max Phil IX-X* (§5.2).

An analysis of the purely philosophical Russellian sources that Gödel read is in order. We discuss the relevant history (§6.1), consider Gödel's correspondence with and about Russell, showing how large Wittgenstein loomed in Gödel's thinking (§6.2), and, finally, we turn to the *Max Phil* notebooks IX-X that Gödel kept while writing (1944), in order to justify our interpretation and set forth conjectures for future scholars to investigate (§6.3).

## 4. The MRTJ

When Gödel proposed “concept theory” at the end of his (1944), he was revisiting the issue of the nature of truth in terms inspired by just the section of *Principia* we have argued stimulated him on the way to incompleteness in 1930. In what follows we examine Russell's introduction of the “multiple relation theory of judgment” (MRTJ) in the *Principia*, and show how this passage gave rise to a host of questions about belief, consciousness, perception, and reflection in Russell's writings after *Principia*.

### 4.1 The MRTJ in *Principia*

Russell first treated the MRTJ in “On the Nature of Truth” (1906/7), though in this discussion he only pointed toward the possibility of the theory, noting difficulties it would face. He first endorsed the MRTJ in 1910. In *Principia* it is presented as a way to argue for eliminating the notion of proposition as “a false abstraction” (*Principia*: p. 44, quoted below). This argument occurs just after the passage concerning *n*th truth quoted in §1. Gödel must have had more than nodding acquaintance with this passage, if our account in §§1-2 is correct, and as we have noted, he referred to it in ((1944): p. 143). Let us examine the remainder of this passage in some detail.

In *Principia* the MRTJ is to provide, as the basis for the (informal) iterated presentation of *n*th truth, “a definition of the simplest kind of *truth* and *falsehood*”. By this Russell makes very clear that he means, not simply judgments caused as a result of perceptual experience in its ordinary sense, but judgments that “are of the same form as judgments as perception” in that “their subjects are always particular and definite” (p. 44). Thus the account will not handle more complex judgment forms, such as “all men are mortal”, nor is it intended to. The base case is the base case.

Russell presented this primordial account, to set up the truth definition, as follows (p. 43):

The universe consists of objects having various qualities and standing in various relations. Some of the objects which occur in the universe are complex. When an object is complex, it consists of interrelated parts. Let us consider a complex object composed of two parts *a* and *b* standing to each other in the relation *R*. The complex object ‘*a*-in-the-relation-*R*-to-*b*’ may be capable of being perceived; when perceived, it is perceived as one object. Attention may show that it is complex; we then judge that *a* and *b* stand in the relation *R*. Such a judgment, being derived from perception by mere attention, may be called a ‘judgment of perception’.

“Mere attention” is the means by which perceptual experience *of* is transformed into a judgment of perception *that*. It “shows” (*Russell's* word!) truth—or, perhaps better, truth shows itself *in* our judgments of perception. The idea—reworked in much of Russell's subsequent writing, and in (1940) in particular—is that actual experience

or “awareness” as a field confronts us immediately, an amorphous field (including particular concepts) to which our minds are in “dual” relations of perception.<sup>52</sup>

The advance to singular “judgments of perception” and the attendant applicability of logic and truth involve an *analysis* of experience: selectivity, attention, and rendering of singled-out elements into the structure of a judgment that is true or that is false. Experience as such does not, according to Russell, contain within itself propositional structure that we can perceive (in a dual mind-object relation). There are no propositions to *be* perceived. Instead, a belief or judgment about certain entities comprises these entities as constituents of the judgment-complex itself, which is also a complex object. Hence, the relation is “multiple”, i.e., of more -arity than dual. And “correspondence” is containment within a conceptual complex.

One of the most significant consequences of the MRTJ, for Russell, was that it allowed him to reject his pre-1910 conception of truth and falsehood as simple (indefinable) properties of complexes that are propositions, a view he had originally taken from Moore (1899). At the end of a series of articles responding to Meinong, Russell made a remark (1904: p. 523) that became well-known enough that Schlick himself reproduced it in his essay “The Nature of Truth in Modern Logic” (1910: I §4)—an essay that, incidentally, it is not wholly unlikely that Gödel knew. Wrote Russell (quoted by Schlick),

It may be said—and this is, I believe, the correct view—that there is no problem at all in truth and falsehood; that some propositions are true and some false, just as some roses are red and some white; that belief is a certain attitude towards propositions, which is called knowledge when they are true, error when they are false.

So that (p. 524) “What is truth, and what falsehood, we must merely apprehend, for both seem incapable of analysis”.

The difficulties with this Moore-Russell view are several, and Russell soon came to see them. As Schlick himself pointed out (1910: I §4), Russell’s view “does not succeed ... in positively defining or making intelligible the real nature of truth; it falls, however, into far greater difficulties still, when obliged to explain how truth is actually *known*”. In fact, as the distinction between true and false propositions is ontologically basic on this Moore-Russell view, there is nothing much more than can be said about truth and falsehood.

With the MRTJ in *Principia* Russell took on board the idea of truth as correspondence in judgment, giving himself room to give a positive account of the nature of truth. He also rid himself of those troublesome, paradox-inducing entities, propositions. On the earlier, non-multiple view, truth and falsehood were taken to be properties on a par with one another, one of each belonging intrinsically to every proposition, or conceptual complex. Belief, or judgment, was taken to be constituted by *dual* relations, the mind bearing a two-place immediate relation of what Russell called “acquaintance” to all the objects of a thought and assembling its grasp of the thought from knowledge of its constituents, which Russell took to be externally related to one another. The *Principia*’s MRTJ presents a sophisticated modification of that

---

52 Russell used the notions of “attention”, “awareness” and “noticing” throughout (1940). *Max Phil* IX [30-31] discusses Russell’s treatment of the self and persons in terms of “awareness” families of experience.

earlier, dual view, and a definite advance in Russell’s appreciation of the complexity of the nature of truth. The MRTJ is a correspondence theory of truth, but, strikingly enough, not one in which a medium of representation (such as language) figures at all.

In *Principia* Russell does not develop his epistemology to fit the MRTJ. Instead he makes a very general appeal to the logical structure of what he calls “judgments of perception”, an appeal we shall scrutinize in a moment (§4.2). But Gödel was aware that Russell’s epistemology had been pursued in a series of more purely philosophical writings on the nature of truth after *Principia*.

The basic idea—having surrendered the notion of proposition—was to analyze a judgement (or “propositional attitude”) complex, not as a relation between a judger and a proposition, but into another complex containing the entities and relations concerned, e.g.,

Othello believes that Desdemona loves Cassio

is analyzed as the following judgment-complex:

$J(O, D, C, x \text{ loves } y)$

And as Gödel wrote (*Max Phil IX [66]*),

*Remark (Philosophy)*: The *Russellian* expression that “propositions” as individual objects are false abstractions would mean that truth is not a class, but a relation concept.

According to the MRTJ’s correspondence account of truth, in a true judgment the subordinate or (more literally) the *conceptually contained* complex of actually exists a representational or correspondence theory of truth, but not one in which the medium of representation, or language, is contained as a constituent. Truth is primarily therefore a property of correspondence (or actual existence) *with* beliefs or judgments, and only secondarily a property of sentences.<sup>53</sup> Moreover, by means of the theory of descriptions Russell’s MRTJ eliminated, not only propositions, but false objectives, the “somethings” that turn out to be false when judgers err.

As Gödel wrote (*Max Phil IX [16-17]*):

*Remark (Philosophy)*: Russell claims: the sentence “A believes bRc” doesn’t have the structure A believes (bRc), but b(R<sub>A</sub>)c. What however re-enters in here (in relation to all simple sentences) appears to be, not a new form, but a new content. All sentences revert to the same form (therefore belief), but with another indication [*Bedeutung*]. Proof of this opinion for example: indirect speech is in the conjunctive, but Latin has another form (*Acc<usativus>c<um>inf<initivo>*).

Gödel frames the MRTJ here, but in such a way that he goes against Russell’s formulation in (1918: IV) that with judgment complexes we simply “have a new Beast (i.e., new form) in our Zoo”. Like the young Wittgenstein, he saw that Russell was trying to throw a new form at the problem, instead of reconfiguring his conception of what a proposition (or concept) is. Interestingly—and such happens quite often throughout *Max Phil IX-X*—Gödel offered a “proof” of this *via* a “grammatical

53 As Russell wrote (1912: ch. XII):

... In fact, truth and falsehood are properties of beliefs and statements: hence a world of mere matter, since it would contain no beliefs or statements, would also contain no truth or falsehood.

remark”: in Latin one cannot formulate indirect speech by quotation of a sentence. In English (or German), where in general indirect statements are introduced by a subordinating conjunction “that”, this is possible, but this only leads to an illusion that what is at issue is the symbolism of sentences. Instead, it is possible, as in Latin, to reconstrue belief claims in terms of a subject accusative + an infinitive, as in “I believe him to be a lover of Desdemona”.

This coheres with a logical assumption Russell made in framing the MRTJ: every verb or relation (e.g.,  $x$  loves  $y$ ) may become the subject of a sentence (e.g., “Loves is a relation”), but no individual (“Desdemona”) may serve as a relation. Gödel is here suggesting that in this particular case, at least in connection with the simplest kind of truth and falsehood, something conceptually (propositionally) new is said when it is said that a something is believed—it is not merely that a substitutional switching of items has occurred within a fixed form (e.g., of a dual relation such as “ $x$  loves  $y$ ”) or that a new form may be added to the complex to explain its relation to the constituents of the belief.

Gödel would have regarded this as a problem of definability. How is one to analyze the judgment—complex in such a way that one avoids postulating a “false objective”—a negative fact that exists if the judgment is false—yet still determines in the analysis the unique thing said by a belief statement, which said thing plays the role of truth-maker if the judgment is true? If one is not to develop a notion of a “possibly false judgment”—as Russell wished not to do—one has little room to manoeuvre if one does not wish to assume anything about conceptual order among items to be given at the outset.

Gödel knew that Russell had explored a variety of options to solve these problems over more than two decades. In (1910b) the mode of combining of constituents was shouldered by the “subordinate” relation among them (i.e., “loves” in “Othello believes that Cassio loves Desdemona”), but this simply adverted to the original Moore–Russell theory, on which the distinction between true and false complexes is ultimate. In (1912: ch. XII), the act of judging did that work.<sup>54</sup>

The full-blown assault on a systematic theory of knowledge for *Principia* based on “acquaintance” was the aim of Russell’s manuscript *The Theory of Knowledge* (1913), the first parts of which were published in a series of five essays in *The Monist* 1914–1915, last part of which Russell abandoned in the spring of 1913 while engaged in lively discussions with Wittgenstein.<sup>55</sup>

Gödel (1944, p. 128) labels as “apparently obvious” the principle “that the signification of a composite expression, containing constituents which have themselves a signification, depends only on the signification of these constituents (not on the manner in which this signification is expressed)”, remarking (p. 130, n. 8) that while Russell “made no explicit statement” of this principle, “it seems it would hold for the logical system of *Principia*, though perhaps more or less vacuously”.

<sup>54</sup> The introduction to (Zalabardo 2012) gives an overview of the versions of the MRTJ; cf. (Ricketts 1996) for a discussion of the problems Russell faced in relation to Wittgenstein’s objections that makes the problem of order central.

<sup>55</sup> See the Editors’ Introduction to Russell (1913) and Pears (1977), Ricketts (1996), Carey (2007), Landini (2007), Potter (2009), and Levine (2013) for general discussions of the mutual impact of Russell and Wittgenstein upon one another in the spring of 1913.

Yet because of his interest in epistemology, during 1903–1913 Russell did not regard the “axiom” as “obvious”, and its status was certainly evolving through his period of struggle with the MRTJ, as Gödel saw. What Gödel meant in having said that it held “vacuously” in *Principia* is interesting to ponder, and would require an account. (Our best guess is that this has something to do, in Gödel’s mind, with the aspect of formalization and proof, which are compositional, rather than with the nature of truth.)

Be this as it may, as to metaphysics, at least in 1942–1943, Gödel took what he called “concepts” to be *structured* complexes, composed of objects appearing “as” playing this or that particular role.<sup>56</sup> This is a view closer to Frege’s, in which the referent of a word is grasped only in the context of analyzing the logical interrelations of thoughts,<sup>57</sup> although Gödel appears to have sided with Russell in suggesting that the only possible subjects of judgments are concepts (*Max Phil* IX [37–8]). Gödel’s treatment of concepts is, also, however, reminiscent of Wittgenstein’s “picture” idea in the *Tractatus*, wherein the content believed is treated as an internally articulated picture, model, or structure [*Bild*].

Both during his salad days with *Principia* and later on in 1942–1943, Gödel had learned from his reading of the *Tractatus*. He had also familiarized himself by 1932 with a goodly portion of Russell’s writings on the theory of knowledge (cf. §6.1 below): (1912), (1921), (1927a) and by at least 1942–3, with Russell (1918) and (1940). As he would have appreciated, the *Tractatus* incorporated more than a few aspects of Russell’s ideas, though it adapted the *Principia* treatment of types and orders to a new framework in which the possible expression of a thought in some language or other was emphasized. For Gödel, it was consideration of the possible expression of a thought in language that had led Russell, Wittgenstein and Carnap to overestimate the significance of the finitary aspects of the expression of thought in language and to underestimate the potential of the MRTJ as a theory of truth for logic and mathematics.<sup>58</sup>

As Gödel would have grasped, however, Wittgenstein developed a view of objects as nodes of possible articulation, thereby bringing modality back into the treatment of the proposition as a structure. The possible–truth–or–falsehood of a proposition (its “*Sinn*”, in the language of the *Tractatus*, its “duality”, as Russell called it (1918: I)), was taken to be internal to the very structure of thought in expressing something that is the case, situated in logical space. And the notion of proposition was reduced to the idea of a possible projection from one space to another, combinatorial possibilities

56 *Max Phil* X [20]:

*Remark (Philosophy):* The distinction between a heap and a structure is that in the structure the parts are “bound” through the fact and these bindings themselves belong to the structure. (Another conception is that it [the structure] is a heap of heaps, but in these the elements are bound through the subordinate part relation [*Teilrelation*]. What is more, a structure somehow has an abstract character (this means: the manner of the elements and the manner of the combination are prescinded from, except for the subordinate part-relation.)

57 Here we follow a suggestion made in (Shieh 2000). Cf. *Max Phil* IX [84] for a remark of Gödel’s about his exploration seen as a kind of “Fregean type theory”.

58 Landini (2007) outlines many of Wittgenstein’s debts to Russell, though without exploring, as Gödel did, what is distinctively new in the *Tractatus* apart from what Landini calls a “doctrine of showing”.

for combinations among objects being incorporated into the very idea of a distinction between truth and falsehood.

Thus was a strong distinction drawn by Wittgenstein (as not by Frege or the earlier Russell) between the way in which a name refers, and the way in which a proposition says just what is the case—and so agreed Russell ever after. But Wittgenstein went a step beyond not only Frege, but Russell too when he construed the “reversibility” of any proposition—its being subject to the operation of negation—as an internal feature of the proposition, not a further conceptual or functional property of a thought or sentence. The *Tractatus* holds that to say that a proposition is true is just the same as to project it, as a proposition, through a sentential expression. The proposition *shows* its sense, i.e., shows how things stand, *if* it is true, and *says that* they do so stand (1921: 4.021).

Here there was a strong Gödelian departure from Wittgenstein, echoing Russell once again. In *Max Phil* IX [37-8], apparently stimulated by Russell’s discussion of assertion in (1940), Gödel suggested that the only possible subjects of judgments are concepts. He is critical of the idea that an “assertion” by means of “the fact that ...” signifies the same thing as the claim itself. This idea—broached in Frege, expressed in the *Tractatus* idea that a proposition shows its sense, and later elaborated by Ramsey<sup>59</sup>—is, Gödel wrote (boldfacing), “**formally false, because** not structurally identical and is therefore also not secured as identical in meaning”. No “redundancy” view of truth was palatable to Gödel, just as no such view was ever palatable to Russell. Both believed in the analysis of truth as correspondence.

We shall return to the distinction between the problem of truth and falsehood in what follows. It rears its head, inevitably, getting first truth into view, and then in the step from first to second truth, where negation enters. That is why it is discussed by Russell in every one of his post-1910 discussions of the formal treatment of truth and falsehood. For Gödel, it will also rear its head finally, in the end, if one takes all the levels in. For now, we merely note that although Russell knew that modality was an issue, at least in 1910-1913 he regarded truth as absolute in the sense of not being subject to modes, even if it must be split into a hierarchy of orders as it is voiced.<sup>60</sup>

Despite this particular issue about modality in Russell’s writings—one that, by and by, would very naturally lead Gödel back to Leibniz—in *Principia* Russell was very clear that the MRTJ should be endorsed so as to define truth in terms of *actual* correspondence. In fact, as we have seen the MRTJ is specifically designed (with the help of the theory of descriptions) to rid us of those possible facts to which false propositions correspond on any dual theory, the (intensional) *something* that one—at least apparently—falsely believes in when one believes a particular thing in error.<sup>61</sup> Of Leibniz’s idea of total harmony or coherence as a mark of actual reality (i.e., the true, the existent) in this best of all possible worlds, Russell stated that “a consistent fairy-tale is a different thing from truth, however elaborate it may be” (1921:

59 As Ramsey wrote (1927: p. 157), following Wittgenstein, “there is really no separate problem of truth but merely a linguistic muddle”.

60 Cf. Landini (2007: app. B); Shieh (2013), Shieh (manuscript), offer lucid analysis of Russell (1905).

61 Hardy, like Gödel, came close to saying that mathematical truths are glimpsed, pointed to, or seen. Unlike Gödel in 1942-3, Hardy (1929: p. 23) professed that he had always “detested” the MRTJ since he felt sure that *something* must correspond to a false thought.

ch. XIII).<sup>62</sup> Correspondence, Russell’s MRTJ, is designed to defeat the idea that all is a mere *façon de parler*.

In (1944: p. 129, n. 7), praising Russell’s correspondence theory of true belief, Gödel duly noted that one

...should expect to be in Russell’s theory a possible fact (or rather the possibility of a fact), which would exist also in the case of a false proposition. But Russell, as he says, could never believe that such “curious shadowy” things really exist.<sup>63</sup>

Given Gödel’s subsequent study of Husserl, Leibniz, and the notion of an intensional object given in perception, it is worth emphasizing that, when it came to “*the* fundamental question [of philosophy], namely, the nature of truth”—Gödel sided with Russell, and not with Husserl, on the issue of correspondence for the basic statements of a theory (we shall return to this point below in §5.1 and §6.3). The question of the absence of false objectives was murkier for Gödel, especially given his embrace of the notion of possibility and his exploration of the analogy between sense-perception and our grasp of simple (indefinable) concepts. Indeed, this murkiness, alongside his uncertainty about how best to develop Russell’s analogy between grasp of simple concepts and sense-perception, may well be sufficient to explain Gödel’s increasing involvement with Husserl over time.

But for us the question is, in what sense did Gödel embrace a “correspondence” view of truth in (1944)? It seems, minimally, that Gödel knew enough to at least try to get beyond the simpler, more easily conventionalized correspondence theory he knew had been developed by, e.g., Schlick (cf. §6.1 below). Our suggestion is that he followed Russell in attaching truth to a multiple relation, or complex, in which correspondence is obtained.

## 4.2 “Judgments of perception”

Let us turn next to Russell’s way of analyzing “judgments of perception” in his *Principia* presentation of the MRTJ. This occurs in the continuation of our above-quoted passage (p. 43):

This judgment of perception, considered as an actual occurrence, is a relation in four terms, namely *a* and *b* and *R* and the percipient. The perception, on the contrary, is a relation of two terms, namely “*a*-in-the-relation-*R*-to-*b*”, and the percipient. Such an object of perception cannot be nothing, we cannot perceive “*a*-in-the-relation-*R*-to-*b*” unless *a* is in the relation *R* to *b*. Hence a judgment of perception, according to the above definition, must be true. This does not mean that, in a judgment which appears to us to be one of perception, we are sure of not being in error, since we may err in thinking that our judgment has really been derived merely by analysis of what was perceived. But if our judgment has been so derived, it must be true.

62 In this same chapter XIII of 1921 on “Truth and Falsehood”, Russell stated explicitly that on the question of truth, Leibniz’s conception of many possible worlds “seems to accord much better with modern logic and with the practical empiricism which is now universal”. He objected, however, that “coherence is not sufficient as a definition of truth”.

63 The reference is to Russell (1918: IV §1).

In the context of an effort to fix the very idea of a truth condition, Russell treats the “see” of perceptual judgment as a “success” verb, comprising actual, true belief.<sup>64</sup> As Russell wrote (*Principia*: p. 43, quoted above), the complex is “capable of being perceived”. It is, in other words, it is there to *be* truly judged. There is then no such thing as a false judgment of perception. For there is no such thing as *perceiving that p* if *p* is not the case.

Continuing our *Principia* quotation (p. 44):

It follows from the above theory that a “proposition”, in the sense in which a proposition is supposed to be the object of a judgment, is a false abstract, because a judgment has several objects, not one. It is the severalness of the objects in judgment (as opposed to perception) which has led people to speak of thought as “discursive”, though they do not appear to have realized clearly what was meant by this epithet.

In *Principia*, then, right at the point where—we have argued—Gödel found his inspiration for applying the Richard paradox in concocting the incompleteness proof, Russell presents at least the beginnings of an account of discursiveness, construing truth as correspondence with a judgment-complex.

On Russell’s (here, merely incipient) account, error arises through misbegotten analysis—although how we are to recognize this certainly cannot be explained by an appeal to anything but a notion undefinable in general, namely, truth, construed here in terms of his MRTJ’s correspondence.<sup>65</sup> The whole treatment certainly does cry out for the development of a theory of knowledge, and this is precisely what Russell aimed to develop after 1910. Lacking such a theory in *Principia*, so far as the foundations of mathematics and logic went, Russell was left to assume, in 1910, that logical considerations (the search for derivations, premises, axioms, and so on) would play the role of an epistemological criterion.

What mattered, however, for Russell’s account of the nature of truth in *Principia* was simply this: particular “judgments of perception”, each and every one being either true or false, *are*. Correspondence, then, *is* or it is not, whether all, some, or none of our judgments turn out to be erroneous or not. That is the fundamental move made with regard to “*the* fundamental question of philosophy”, the nature of truth.

Russell proposed a *definition* in *Principia* (p. 43):

In fact, we may define truth, where such judgments are concerned, as consisting in the fact that there is a complex corresponding to the discursive thought which is the judgment. That is, when we judge ‘*a*-in-the-relation-*R*-to-*b*,’ our judgment is said to be true when there is a complex ‘*a*-in-the-relation-*R*-to-*b*,’ and said to be false when

64 This analysis is usually attributed to Austin, as in Putnam (1999), and more recently Travis (2013). Of course, Austin and Wittgenstein offer powerful arguments against the kind of sense-data and solitary language treatments that Russell would go on to elaborate in (1940). But, perhaps surprisingly, Russell’s *Principia* analysis already took a first step in the direction of Austin. Indeed, Wittgenstein accepted Russell’s rejection of the dual account of perception from the beginning, on the basis of working through these very same passages in *Principia*.

65 Russell went so far as to write, much later on (1940: ch. XV), that “the sole method of discovering error is, I believe, the experience of surprise owing to a disappointed expectation”. This remark, and the numerous discussions of surprise in Russell (1940), should be compared with Wittgenstein’s “Appendix on the Surprising” in (1978: I App. II), which is in part a response to Russell’s remarks about “expectation” and finitism in (1936), with its roots in this very same *Principia* passage.

this is not the case. This is a definition of truth and falsehood in relation to judgments of this kind.

We complete our *Principia* quotation with the mention of “redness” to which Gödel alluded ((1944: p. 142), quoted above):

It will be seen that, according to the above account, a judgment does not have a single object, namely the proposition, but has several interrelated objects. That is to say, the creation which constitutes judgment is not a relation of two terms, namely the judging mind and the proposition, but a relation of several terms, namely the mind and what are called the constituents of the proposition. That is, when we judge (say) ‘this is red’, what occurs is a relation of three terms, the mind, and ‘this,’ and red. On the other hand, when we perceive ‘the redness of this’, there is a relation of two terms, namely the mind and the complex object ‘the redness of this.’ When a judgment occurs, there is a certain complex entity, composed of the mind and the various objects of the judgment. When the judgment is true, in the case of the kind of judgments we have been considering, there is a corresponding complex of the objects of the judgment alone. Falsehood, in regard to our present class of judgments, consists in the absence of a corresponding complex composed of the objects alone.

A consequence of Russell’s treatment of “the simplest kind of *truth* and *falsehood*” is that there is no common objective (no “apparent fact”) shared by an experience of a complex that can give rise to a true judgment of perception that  $p$ , and one that cannot, when  $p$  is not the case. There *is* no judgment of perception that is *of*—i.e., corresponds to—a falsehood, and no perception of anything that-is-false. As those who analyze and verbalize our judgments of perception, *we* are the ones who may go wrong in judgment. Truth, however, is what it is, in terms of what is “capable of” *being* perceived or judged (*Principia*: p. 43). As Russell wrote (1912: ch. XII), “minds do not *create* truth or falsehood”. In a world without beliefs, there would be no occurrences of error or knowledge. Yet the results of judging, judgments or beliefs, such as they are, simply are, and what corresponds in the case of true judgments of perception is there to be judged, truly or falsely.

According to this account of judgments of perception, it is clear that there may well be look-a-like strikings within (dual) perception. This, let it be said, is obvious. To use the example from fiction Russell favored in (1918), Othello’s glimpse of the complex of Desdemona and Cassio entangled his mind in (dual) perception with the constituents that he assembled to form a judgment of perception that was false (that she loved Cassio) whereas he might have assembled them to form one from the same constituents that was true (that she did not). Another kind of (dual) experience (of, e.g., a diagram of a heptagon) may give rise to a judgment (an analysis) that goes wrong (e.g., if we judge that *this* was constructed with ruler and compass in the classical way).<sup>66</sup>

Experience, Russell always believed, is rich, even if not propositional: it is not reducible to what may be verbalized in a single judgment, at least so far as we know.

66 Of course, there must be some common constituent involved, according to Russell’s doctrine of acquaintance, in a particular judgment that  $p$  when  $p$  is the case, and the particular judgment that  $p$  when it is not. Otherwise, the only thing in common are the words used to express the judgment, but they are about different objects. Russell’s “logically proper names”, such as *this*, are designed to ensure direct connection with particulars (including particular universals) and rule out mere verbiage of merely apparent disagreement, which cannot express truth and falsehood, on his view.

He was quite explicit, from 1913 on, that there is no bound to be assigned to its complexity.<sup>67</sup> In (1940) Russell even adduced the infinite to argue—against Carnap and the positivists—that it cannot be the case that every observable fact may be put into words, so that no appeal to a linguistic framework could be said to *constitute* the notion of a fact that corresponds to a true judgment. Just here Russell mentioned the idea that words themselves, as ontological items, may well figure in facts, and he uses this to draw in the infinite. Praising Gödel’s arithmetization of syntax as “the best way” to try to make sense of the totality of “possible” statements that may be made (p. 70f), his general argument concerning experience is this:

Can we say: “he will know everything that can be known by observation alone, but nothing that needs inference”? Let us first alter our question, and ask, not what can he know, but what can he express in words? To begin with: if he can put every observable fact into words, he must have as many words as facts; now some words are among facts; therefore the number of his words must be infinite. This is impossible; consequently there are facts he leaves unexpressed. The case is analogous to Royce’s bottle with a label on which there was a picture of the bottle, including, of course, a picture of the label.<sup>68</sup>

The whole discussion is strongly reminiscent of (and in fact historically derives from) Dedekind’s Theorem 66 in his (1888), where the existence of a “simply infinite system” is proved. Dedekind observed that there must be in the world items that are not thoughts, since the object of a thought differs from the thought of it. Hence, through reflection on his thought of himself, the thought of the thought of himself, the thought of the thought of the thought of himself, and so on, the existence of a Dedekind-infinite system mapped into a subset of itself could be established through the power of thought alone. However, unlike Dedekind Russell has adopted neutral monism, so that there is no (dual) conception of a “self” and its thoughts as objects in the analysis of a judgment complex.

Russell is of course aware of the need for an axiom of infinity in (1940). In the above-quoted passage he emphasized, in a different vein, the potentially unbounded complexity of experience and of conceptualization. As he saw it, conceptualization is not simply read off of anything given directly in experience itself: *analysis* is a prerequisite of judgment. As Russell (1910a: ch. VII) and Frege (1918) both argued, images, strikings, and perceptual (dual) experiences do not come labelled with “T” and “F”, and even if they did, these labels would be of no help. Truth is just not passively given in (dual) perception, it requires the activity of judgment. The important consequence to draw from this with respect to the notion of a truth condition is that this very notion itself is complex, and not a (dual) psychological or perceptual element of experience we can passively register.

Later on, in his retrospective account of his philosophical development, Russell would call his belief in singular “concepts”—and not merely those applying to physical things—an “initial prejudice” with which he began, and “which has been perhaps the most important in all my thinking” (1959: p. 98). As he wrote (p. 136-7),

<sup>67</sup> Russell (1940: p. 336) wrote: “There is no theoretical limit to the complexity of the object of perception or of the structure affirmed in judgments of perception which the object verifies”.

<sup>68</sup> Royce’s example, using the example of maps within maps, may be found in the Supplementary Essay to his Gifford Lectures, First Series, “The One, the Many, and the Infinite”, §III I (1900: p. 503).

In this investigation [of the nature of belief] I tried to proceed from what is most simple and primitive and unquestionable towards the more difficult complex and doubtful cases. I should have thought this procedure the obvious one to adopt on general methodological grounds, but I found that most of the writers who concern themselves with the definition of ‘truth’ proceed in a quite different manner. They start with what is complex or questionable, such as the law of gravitation or the existence of God or quantum theory. They do not trouble their heads with plain matters of fact, such as ‘I feel hot’. This criticism applies not only to pragmatists, but equally to logical positivists. Philosophers of almost every school fail to investigate our knowledge of particular facts, and prefer to start their investigation with our knowledge of general laws. I think this is a fundamental error which vitiates most of their thinking.

Gödel, like Wittgenstein, followed Russell here. But the natural progress of his own mathematical and logical thinking—the building of complexes within complexes containing “correspondence” secured through reflection (through definability)—drove him to do so. Philosophically speaking, the whole needed a point of view in the end, a clarification “in terms of which finally everything is to be interpreted” (1944: p. 142). Success in analyzing the most fundamental problems—here, truth—often comes late in a thinker’s life: witness Dedekind, who had to work his way down to what he called the “naïve point of view”, the “most protracted” part of his labor, as he described it, when he came to analyze the basic foundation in infinite systems of the natural numbers.<sup>69</sup> Both for Russell and for Gödel, the aim was to clarify the simplest distinction of all, that between truth and falsehood, in a simple way.

### 4.3 The analogy with sense-perception: simple notions

At this point we reach the need for simplicity, the basic notions in terms of which the whole theory is to be erected. Gödel’s view was that there *must* be (particular) simple concepts, or else the whole theory of truth as correspondence, the MRTJ, collapses.<sup>70</sup> He also held (shades of the Vienna Circle, as well as the later Wittgenstein) that it follows from the idea that we create concepts by mere stipulation that there are no simple concepts (*Max Phil* IX [56]). Most fundamentally, he saw that the difficulty Russell was facing in *Principia* concerned the very idea of a truth *condition*: he wrote, boldfacing (*Max Phil* X [17]), “**If we could not somehow perceive concepts, we could not understand a single proposition [Satz]**”.

Here we must stress that Gödel only reached for the analogy with sense-perception, at least in 1942–1943, in attempting to elucidate our grasp of the *primitive* undefined notions of a satisfactory analysis. The analogy with perception is an *analogy*, an extension of the verb “to see”, for better or for worse. And here Gödel is quite clear in (1944) that he has been inspired, not only by Russell’s realism but also by the analogy

<sup>69</sup> On this issue of getting back to the “naïve” level at the foundations, cf. Floyd (2013b).

<sup>70</sup> *Max Phil* X [17]:

*Remark. Philosophy.* One perceives concepts insofar as one understands propositions (or their sense [*Sinn*]). One perceives them however as independently existing [in a higher degree than bodies [*Körper*]], where it is provable that sentences about these cannot be back-translated and at the same time <more> immediately than bodies, because [there] clearer perception is apparently excluded.

Wang (1996: p. 232–3) reported conversations with Gödel about perceiving concepts, but without making the restriction to primitive notions explicit.

Russell repeatedly drew between our grasp of primitive notions what the early Russell called the “indefinables” and sense-perception. In any analysis, there is a residue, that which is elucidated, but not justified. The “simple” notions are those where the residue makes itself felt, is explicitly and directly displayed in an immediate (non-inferentially rendered) way.<sup>71</sup>

Alluding to “one of [Russell’s] earlier writings”, probably the section of *Principia* where Russell had given his reasons for accepting the Axiom of Reducibility (p. 59), Gödel wrote (1944: p. 127):

... the axioms [of logic and mathematics] need not necessarily be evident in themselves, but rather their justification lies (exactly as in physics) in the fact that they make it possible for these “sense perceptions” to be deduced; which of course would not exclude that they also have a kind of intrinsic plausibility similar to that in physics.

Gödel is not primarily thinking here of Russell’s external world program, in which a top-down, ancestral-like construction would be used to construct ordinary objects out of collections of possible perspectives in space (sense data), although the analogical remarks here in (1944) cover this case as well, and there is some reason to think that the *Principia* version of the MRTJ that Russell had in mind would have involved an analysis of true judgments in terms of equivalence classes of belief-facts.<sup>72</sup>

Instead, Gödel was thinking of Russell’s analogy between our grasp of the indefinables (e.g., truth) and sense-perception, our way of knowing the simple notions involved in a correct analysis.<sup>73</sup> After all, if one looks toward the undefined notions in a theory treated as foundational, one is struck by several features of their role: their apparent lucidity and transparency in being isolated within the axiomatization; their particularity; their directness, coupled with a hazy status of being perhaps further analyzable, until the judgment renders what is seen clear, as in a sharp outline; our non-deductive, non-demonstrative, and seemingly immediate grasp of them; their fertility in being articulated and ramified in the context of an axiomatic theory.

These features of the undefined notions drive one very naturally toward the sense-perception metaphor, as one can see in the case of the attempt to describe the content and character of our own perceptual experiences as they figure in singular perceptual judgments, true or false. The visual field is indefinite in extent, and may be “cut” across a seemingly unbounded number of differing dimensions, through predication. Yet whenever we do select for judgment a particular predication, we must leave something undefined.

In this vein Russell had famously written (1903: p. xv):

The discussion of indefinables—which forms the chief part of philosophical logic—is the endeavor to see clearly, and to make others see clearly, the entities concerned, in order that the mind may have that kind of acquaintance with them which it has with

71 Wang (1996: p. 300) wrote that Gödel believed “we arrive at all our primitive concepts by idealization”, but it is fairly clear that Gödel conceived of “idea” in something like Kant’s sense of an undefined totality or limit toward which we strive, but which it does not make sense to say that we know in any empirical way.

72 Linsky (1999: ch. 7.2).

73 Wang (1996: p. 245) surmises a specific reference for the above quoted remark by Gödel (1944: p. 127), an essay Russell first published in French (1906a). Whether or not this is so, the broader analogy may be found in many writings of Russell that Gödel knew.

redness or the taste of a pineapple. Where, as in the present case, the indefinables are obtained primarily as the necessary residue in a process of analysis, it is often easier to know that there must be such entities than actually to perceive them; there is a process analogous to that which results in the discovery of Neptune, with the difference that the final stage—the search with a mental telescope for the entity which has been inferred—is often the most difficult part of the undertaking. In the case of classes, I must confess, I have failed to perceive any concept fulfilling the conditions requisite for the notion of *class*.

It is this part of the analogy between knowledge of abstract notions and sense-perception that forms the *locus* for Gödel's realism. Gödel does not simply embrace and defend an ontology or conception of truth separate from the human mind and/or linguistic convention, or postulate intellectual faculties for knowing, or insist on inflicting the law of the excluded middle willy nilly across all mathematical sentences. Nor did he assert naively, without further reflection, that we can “perceive” concepts. He insisted rather on the centrality and importance to logic of discussion and elucidation of fundamental notions, especially proof, definability and truth.

In (1944) Gödel did not rely on a theory of intuition.<sup>74</sup> He did, however (p. 151) speak of perception in connection with the fundamental concepts, writing that “the difficulty is only that we don't perceive the concepts of ‘concept’ and of ‘class’ with sufficient distinctness, as is shown by the paradoxes”. Gödel wrote explicitly, though tentatively (p. 127), that the “logical fictions” of Russell—apparently referred to by singular phrases that are nevertheless eliminated as incomplete symbols by the theory of descriptions—“need not necessarily mean (according to the sense in which Russell uses this term) that these things do not exist, but only that we have no direct perception of them”.<sup>75</sup> In other words, to be defined is *not* to fall under the auspices of the analogy with sense-perception.<sup>76</sup>

At least in 1942–1943, we take Gödel to be limiting the analogy with sense-perception to the indefinables, and then following Russell in rejecting a sharp distinction between concepts and objects. He is then simultaneously taking on board a kind of “unsaturatedness” with respect to those concepts whose recognition is said to be gained through something like sense-perception. Gödel does not subscribe to a sharp concept/object distinction, as Frege did. Again, Gödel follows Russell here. And like Russell in (1903), Gödel does not claim to have “perceive[d] any concept

74 Parsons (2010, p. 335).

75 Russell wrote (1959: p. 101), “Nominal entities of this sort may or may not exist, but there is no good ground for assuming that they do”.

76 It is interesting that Gödel did not use the analogy with perception when he wrote in 1964 (1985: pp. 369–71) that Turing had given a “precise unquestionably adequate definition of the general concept of formal system”. However, Wang (1996: pp. 232–233) reports him using it in conversation about Turing's analysis. This may be due to Gödel's concerns in the philosophy of mind, involving the possibility of a Leibnizean view of human experience as infinitely rich (see [Webb 1990]) for a penetrating analysis of Gödel's remark): he appears to have been convinced that Turing was committed to a discrete, finitistic conception of experience. This would fit with the interpretation of Gödel in 1942–1943 that we suggest here. Floyd (2012a,b, 2013a) argue that Wittgenstein understood the vividness of Turing's analysis differently, without any particular commitment to a philosophy of mind, while taking the concept of a human being “following a rule” to be a primitive notion.

fulfilling the conditions requisite for the notion of *class*.<sup>77</sup> The indefinable notion here, the one Gödel is primarily concerned with, is that of *concept*, which is intended to analyze the Russellian notion of a *propositional function*. For that latter notion, sitting on the razor's edge between a metaphysical entity and a piece of language, requires the notion of truth to buttress our grasp of it, as Russell and Gödel agree.

In the end, Russell's point in the *Principia* passage (p. 43f) is simple. It *has* to be simple, for this is the simplest kind of truth and falsehood, treated from the point of view of logic, and not of psychology. The step to judgment from (dual) perception, even when guided by prior true judgments of perception, may fall to pieces in the multiply-related structure of our judgment: truth may be lost along the way. How could it be otherwise? Experience, whether perceptual or mathematical, is rich and indefinite in extent. There are many different judgments of perception, many different conceptualizations, to make of any one. And experience does not come labelled with conceptual structure that is visible to the (dual) eye. But, conversely, in a judgment of perception (in Russell's sense), we judge a particular conceptual complex to be true of *one particular thing*. Russell's treatment of "the simplest kind of *truth* and *falsehood*" is designed to capture a very simple, naïve idea: our particular judgments may be true, or they may be false, and, when true, they are made true by an existing fact.

#### 4.4 Russell's theory of knowledge

Large difficulties confronted Russell, in the work he wrote after *Principia*, in developing a theory of knowledge. He acknowledged this and spent decades attempting to fashion a view. He had to work out how to make sense of negative judgments of perception ("this is not white"), as well as the issue of perceptual errors shared by different viewers. Above all, he sought a perspective from which both logic and epistemology could find their proper role.

In conversation with Wittgenstein 1912-1913, while writing (1913), Russell made a clear step forward, cleaning up the *Principia* by restricting the base case to what we would now call "elementary propositions", those containing neither quantifiers nor truth-functional connectives, including negation.<sup>78</sup> Yet Russell's struggles with the nature of truth continued. As late as (1940) he found "rather difficult problems" involved in stating the difference between negation and falsehood at the simplest point *via* the transition from perception to judgment (p. 81ff).

<sup>77</sup> The notion of *set* might be said to be defined through its axioms, hence simple. In *Max Phil IX* the analogy with sense-perception was used by Gödel in connection with this notion, when he discussed the most basic mathematical notions (set, identity, number). Gödel held [90-91] that "mathematical unities are the simplest objects that one sees precisely" and that they have "nearly only negative properties" in the sense that ordinary empirical concepts (Gödel's example is the concept of *raven*) don't apply to them. A further discussion of the notion of *set* lies outside the scope of our essay. Our point here, moreover, is that that notion is treated alongside many others in *Max Phil IX-X*, because here Gödel was concerned to treat the notion of truth, not primarily the axioms of set theory.

<sup>78</sup> Levine (2013) explains that in the first edition of *Principia*, Russell had not yet separated what we would call "molecular", unquantified sentences out from the elementary level, arguing, successfully we believe, that this shows a disanalogy between the Tarskian recursive presentation of truth alleged for *Principia* by Landini (2007). Our purely philosophical comparison with Russell (1940) equally illuminates, through *Max Phil IX-X*, that Gödel was also not regarding Tarski as providing the kind of analysis needed for "the simplest kind of *truth* and *falsehood*".

Gödel followed Russell, even in 1942-1943, in supposing that judgment (belief), which is capable of being true or false, derives from experience of particular concepts or individuals, and consists in conceptual complexes. As a dual relation between mind and object, experience is neither true nor false, but *via* analysis—“noticing”, and “attention”—the possibility of truth comes about in judging. “The simplest kind of *truth* and *falsehood*”, for the simplest propositions, may be defined in terms of correspondence within a conceptual complex. But for epistemology, the connection between dual experience or “perception” of particulars and our advance to judgment and belief must be explained.

Here it is to be emphasized that the notion of “experience” was, for Gödel, loose and very broad, workaday in the mathematician’s sense. It was to include the mathematician’s experience with and knowledge of mathematics, grammar and language (including not only linguistic facts, but also the concepts and proofs so far encountered), as well as experience with history, emotion, physics, life, and human psychology. This explains why *Max Phil* IX-X contain so many remarks on this wide variety of topics.

Moreover, this loose and evolving sense of “experience” is not very far from Russell’s conception, at least insofar as we bracket the issue of sense-data and focus on the general idea of what epistemology may accomplish. Gödel’s reading of Russell’s (1927a) in 1932 would have provided sufficient grounds for understanding epistemology in this sense. There Russell addressed the problem of showing how physics could be *true*, and he construed the problem as having two parts (p. 7): “to assimilate the physical world to the world of perceptions, and to assimilate the world of perceptions to the physical world”. This would be done by means of conceptual articulation—ultimately, the fashioning of descriptions and characterizations to erect the assimilation with sufficient precision that the objects of perception would “interpret” physics in the sense of making its theories actually true. In this context, Russell stressed, the whole enterprise would take place within the confines of “neutral monism”, thereby justifying the doctrine that the observing “self” disappears on analysis, consciousness being derivable from descriptions of relations (p. 10). Notably, this would allow him freer reign in assembling the notion of truth from judgment complexes, the latter now construed as single (conceptual complexes of) events in space-time.

Thus on the kind of view Gödel held—inspired by Russell—knowledge in logic and mathematics involve a kind of *selectivity* within experience. Although this theory cannot provide a general definition of “truth”, it can ensure that content is diffused through the whole hierarchy used to reconstruct mathematics. For, as to epistemology, the use of formalizable proof methods and recursively presented partial notions of truth controls conceptual articulation and tracks truth up the backbone of the ordinals by means of which the universe is constructed. Definability, analysis, truth, proof: these are the rungs of the ladder that are there to climb up, relying on the possibility of unlimited indexing by transfinite ordinals. Those ordinals themselves are better and better analyzed by means of analysis, just as richer and richer conceptual elements may be extracted, in an indefinite and unlimited way, from particular perceptions, and new concepts and definitions devised to group truths. Those concepts may then themselves be added to the realm of “experience”, enriching what may be regarded as the “perceptual” basis.

In (1940) Russell explicitly stated (p. 329) that “in the theory of knowledge, what is fundamental is noticing, not sensation”. Gödel would have seen that after 1918 Russell had surrendered his notion of knowledge by “acquaintance” (this notion does not appear in (1940)), and that he had replaced that notion, at least in the case of experiential knowledge, with the concept of “noticing”—a concept that had been anticipated already in *Principia* with the phrase “mere attention” right at the point where Russell introduced the MRTJ (p. 44). In Gödel’s *Max Phil IX* the word for “awareness” used is [*Aufmerksamkeit*], and he frequently reverted to discussion of whether it is an action or a passion, and in what sense it might allow for truth bearing contents to emerge ([22,26,29,41,50,60,85]; at [30] Gödel even uses the English “awareness”. As in (1944: p. 130) he is perfectly aware of the danger of irrelevant psychologism entering into the discussion.

Like the notion of truth in general, Russell’s (1940) idea of “noticing” remained undefined; unlike truth (at least for Russell) it is always a matter of degree and partial appreciation of experience as a whole ((1940: p. 50); cf. (1959: p. 104)). The concept is ultimately intended to sophisticate Russell’s theory of experience to allow for a richness suitable for the step into metaphysics, but without assuming the independent reality of consciousness.<sup>79</sup> Gödel commended this, but went beyond Russell in regarding that step as a necessity for the analysis of mathematical truth.

In *Max Phil IX* Gödel goes so far as to toy with the idea of truth as a matter of degree, broaching it in connection with his idea that the antinomies could be explained by “singularity” points (cf. 1944: p. 150):

[48] *Remark (Grammatical)*: The final ground for the antinomies is that we don’t see what the  $\in$ -relation really is [in the realm of concepts], but we see a substitute [*Ersatz*] in that which we have constructed. Just so little do we see what the concept “concept” is.

*Remark (Grammatical)*: Because of my analysis of the (*intensional*) antinomies (in the *Notre Dame* lecture) there are (if the propositional calculus is assumed) two possibilities for seeing a mistake: 1.) Not every defined propositional function defines a concept (that is apparently the correct one, for it would have the result that  $x \notin x$  is actually defined for all objects and nevertheless defines no concept). 2.) The concept of “definable concept”. That means, there are “limit values” or “singularity points” to which it is not applicable [note: it could however also mean that there is a set of sharply defined approximations instead: concepts of individuals, for concepts, etc. (this is really (?) the whole essence of type theory)]. It must however also be that the concept of “meaningfully applicable” isn’t defined [otherwise one could define: false everywhere, where senseless [*sinnlos*]]. Type theory makes assumption 2., but belonging to it also corresponds such concepts as “concept” of an infinite set of different concepts with differing realms of sense [*Sinnbereichen*].

<sup>79</sup> For an interesting argument that we do not in general know our present experiences, on pain of an infinite regress and collapse of the distinction between the experience and the knowing, compare Russell (1959: p. 105). Here, as in his earlier MRTJ, Russell is insisting that “there is a sense of ‘knowing’ in which, when you have an experience, there is no difference between the experience and knowing that you have it”. The story of Russell’s own development away from a Moorean emphasis on foundationalist epistemology and a strong principle of acquaintance is complicated; Russell tells us only a certain amount in his (1959). A very useful overview rooting Russell’s development in features of his mathematical practice is Levine (2009).

This line of thought, however enigmatic,<sup>80</sup> makes at least some sense in the context of our Gödelian reading of Russell. Since our description of experience in terms of judgments of perception (in Russell's sense) is always only ever, for all we know, "approximate" and "blurred", taking the step to conceive of experience as infinite would require some such treatment of "singularities", a theory of "limit" points. Leibniz too, as Gödel knew, took conscious perception to be sometimes "blurred": experience could be actually infinitely complex, yet approximated by our conceptual means (as when we hear the roar of a wave, itself composed of a potential infinity of droplets).

Gödel was, we have said, keenly aware of Russell's own evolution, and greatly engaged, in 1942–1943, with Russell's naturalistic turn. Yet he was most unhappy about Wittgenstein's influence on Russell. Let us briefly revisit the picture he had of Russell's development from the perspective of Russell's work on truth before 1918 before turning to Russell's treatment of neutral monism in (1940).

In *Principia* and (1913), "complexes" were assumed by Russell to be composed of particular constituents standing in wholly external relations to one another—this is what Schlick called (1925: I §10) the "ingenious" part of Russell's view: in it truth "is built up entirely out of differences that characterize various kinds of relations".

Yet just this gave Russell the greatest difficulty. For, given that he required a general theory of knowledge, at this point he faced what may be said to be a fundamental problem of how to conceive truth. The mind's selectivity through awareness in a dual relation, rendered in judgment as a conceptual complex, was, he would have said through (1913), to be regarded as ultimately bottoming out in direct, singular "acquaintance" with each of the constituents that the judgment was about. After all, if judgment is itself a fact or complex (one in which correspondence holds, if the judgment is true), then it must be that in a judgment a mind bears the judging relation to the specific complex that would have to be combined if the judgment were true.

The bottoming out in acquaintance was to explain understanding, and also contact between mind and the objects of knowledge. Yet a judgment–fact must, it would seem, involve correspondence or failure of correspondence, not merely with *what* is combined, but also with *how* the constituents are to be combined. Russell had to find a way to account for all possible combinatorial possibilities of substitution in complexes in order to single out from these the precise truth condition associated with each judgment–complex. Yet, at the same time, he believed that logical possibility, as a notion, should be eliminated from logic (1913: 111).<sup>81</sup>

Thus Russell faced what has been called the "mode of combination" problem.<sup>82</sup> It is really a whole nest of problems about how to regard the fixing of a truth condition given Russell's "ingenious" proposal. Moreover, it was on precisely this nest that the young Wittgenstein fastened in his discussions with Russell during 1912–1913.

80 So it is called in Parsons (1990).

81 The importance of this point is emphasized in Shieh (2013) and (Shieh manuscript).

82 Cf. the Introduction to Zalabardo (2012) for an overview of the three versions of the MRTJ Russell developed between 1910 and 1913; cf. also the editors' introduction to Russell (2013). Among literature reconstructing the history and nature of this problem are Pears (1977), Griffin (1985), Hylton (1990), Ricketts (1996, 2001), Linsky (1999), Carey (2007), Pincock (2008), Potter (2009), Sullivan and Johnston (forthcoming).

As we have seen, in (1913) Russell proposed analyzing judgment-complexes such as

Othello believes that Desdemona loves Cassio

as

$J(O, D, C, x \text{ loves } y)$ .

Yet *which* complex precisely was this to be taken to be? First, there is the problem of order: How is the difference between Desdemona loving Cassio and Cassio loving Desdemona to be captured in the analysis?<sup>83</sup> Some relations are asymmetric and some are not, so that some ways of individuating facts appear indifferent to permutation of constituents in a complex in a way that others do not. Second, which relation is responsible for this, the relating relation (of belief or judgment) or the subordinate relation (of loving)?<sup>84</sup> Finally, there are a number of problems about negation, generality, and logical form. It must make sense or be possible that Othello, Desdemona, Cassio and loving all combine in this way, or not. In other words—something to which Gödel, with his focus on the theory of types, would have been sensitive—there must be an account of the type as a constituent of some kind, as well as the way in which the law of excluded middle applies to the complex. But what kind of account might be given?

Wittgenstein pressed many of these questions upon Russell in the spring of 1913, leading him to abandon the manuscript of (1913). Gödel could not have known of the details of their interaction. But he would have likely grasped—having read the *Tractatus*, and Russell (1918, 1940)—that a notion of *order* would have to be somehow added to Russell’s account to make the MRTJ work. As Gödel also would have known, the *Tractatus* idea of a proposition as a picture showing how things stand if it is true was an offer by Wittgenstein to Russell of a conception of structure or order sufficient for the MRTJ. Wittgenstein’s idea turned on taking the notion of a “possibility of projection” of one structure into another as fundamental: the proposition is construed as a projection from a possible medium of expression—itself a fact—to another fact, and truth is said to show itself in what we “say” or express about what is (or is not) the case.

On this account, the notion of “logical possibility” was not reduced or eliminated from the foundations of logic, as Russell had said it should be. Wittgenstein was claiming instead that one need not isolate from an independent point of view or perspective the specific form (possibility) of a false judgment complex, for the possible truth or falsity of a projection would be an internal feature of the proposition itself. Moreover, in the case of a belief statement, the subject or self (e.g., Othello) would disappear on analysis.

In the *Tractatus*, just where he discussed the “old” (dual) theory of propositions of Moore and Russell so as to bring forward Russell’s multiple relation view, Wittgenstein remarked (1921, 5.5421) “there is no such thing as the soul—the subject, etc.—as it is conceived in the superficial psychology of the present day. Indeed a composite soul would not be a soul any longer”. To spell this out, he adduced an ambiguous perceptual

83 Cf. Griffin (1985), Ricketts (2001).

84 Cf. Zalabardo (2012), Introduction.

object, the Necker cube. This drew an important consequence from Russell's rejection of the dual theory of propositions and truth. The ambiguous Gestalt figure shows that different configurations of relations may be projected from the very same empirically given sign configuration or sentence. It follows that truth is not a simple property of sentences, and that judgment is not dual. But then the judging self is neither a formally nor an experientially given whole or constituent.

The judgment-fact, according to the *Tractatus*, must be understood to be a multiply-complex projection analyzed in a logical space of possibility. The modification of the *Principia's* "judgments of perception" was this: although we cannot (in a dual way) "see" a possibility in the same way that we can "see" a complex, we may in the activity of judging express what is true or false, so long as we take such expression to be framed by a "scaffolding" of logical possibility, "selected", as it were, as a specific place within the backdrop of logical space.

The price Wittgenstein asked Russell to pay was to make the primordially and unity of truth unsayable. As Gödel knew, in his Introduction to the *Tractatus* Russell rejected Wittgenstein's offer. Russell suggested instead the idea of a hierarchy of languages, thus moving to make explicit the implicitly typed logical architecture of *Principia*, and the systematic ambiguity of truth. And in the second edition of *Principia*, a constructivistic theory of orders is proposed.

Thus—from Gödel's point of view, crucially—Wittgenstein had stimulated Russell to develop the idea of a hierarchy of levels of truth, an essential step on the road to incompleteness. And yet—from Gödel's point of view, disastrously—Wittgenstein had led Russell to surrender his conception of truth in logic and mathematics, labeling them "tautologous", mere figments of symbolism. For Gödel, this placed exactly the wrong weight on the expression of thought in language, and misconstrued what the theory of logical symbolism and the notion of truth require. Gödel resisted altogether Wittgenstein's emphatic unwillingness to spell out further traits of the notion of truth. He regarded this move—and not without some justification, given the effects of the *Tractatus* on Russell and Carnap—as leading inevitably to a collapse of the notions of logical and mathematical truth into mere properties of language.

Of course, unbeknownst to Gödel, by 1934 Wittgenstein too had rejected the *Tractatus* idea of a "possible projection" in logical space, and refashioned the idea against the more anthropological backdrop of "language games" and "forms of life". But, as is clear in *Max Phil IX-X*, it is logical, and not anthropological ideas of meaning that interested Gödel. As he wrote (*Max Phil IX* [13]),

*Remark (Philosophy)*: 1. The "concepts" in an objective sense suffice probably for a weakened *Thesis of Extensionality*, in which e.g. only one concept exists which accords with all similarly powered sets. That is the apparent form of the fact that there are many fewer concepts in the objective than in the *anthropological* sense (another is that for many objectives nothing *anthropological* exists); for example, in general for  $p \vee q$  no "objective fact [*Sachverhalt*]" is, rather merely something brought into the world by us (because of our uncertainty), namely exactly when it isn't a "consistent aspect" of the world. In this sense are the forms [or: formulas] and the formations something subjective. (*Compare* also [the] contradictory concept  $x \notin x$  [nothing exists].)

Gödel's idea of concept theory in *Max Phil IX-X* is intended to adapt the idea of possible projections to the setting of the first edition of *Principia*. He is suggesting that Russell might retain the MRTJ by taking the ordinals' structure as basic, and

then, using them to index the ascent to higher types, overcome the linguification of logical and mathematical truth Russell had come to embrace under Wittgenstein's influence. The idea of truth as correspondence would be controlled and made explicit by means of definability. And axiomatic proof, through the arithmetization of syntax, would allow for a sufficiently rich conception of epistemology to suit the MRTJ's needs. The basis would still be in "judgments of perception", but these selectivities, taken at the outset, could be revisited and revised, by analysis, in the ascent up the (ordered) hierarchy.

## 5. Russell's *Inquiry into Meaning and Truth* (1940)

*Max Phil* IX-X show Gödel grappling with the evolution of Russell's thought, comparing and contrasting his earlier (*Principia*) and later (1940) presentations of the nature of truth, and the MRTJ in particular—including its more purely philosophical aspects. These, as we have seen, involve a theory of belief, consciousness, self, awareness, experience, and noticing. This terrain in Russell was relatively uncharted territory when Gödel entered it. Yet it stands to reason that as he was writing (1944), given his own penchant for historical research, Gödel would have felt it incumbent to read at least the then-latest Russell, as well as the earlier writings that had so inspired him initially (on what these were, see §6.1).

The overarching features of Russell's handling of the formal treatment of "the simplest kind of *truth* and *falsehood*" in (1940) do not deviate from the first edition of *Principia* with regard to "judgments of perception". However, as in other post-*Principia* texts, Russell explicitly discussed Wittgenstein's impact on him. And here, with respect to epistemology, Russell's whole treatment of the problem is framed by his naturalistic turn toward psychology, initiated in his (1919b).<sup>85</sup> Whereas in (1913) Russell attacked "neutral monism" explicitly, so as to defend his appeal to "acquaintance" (an act-object model of conscious awareness), in his books written after his psychologistic turn (e.g., Russell (1921, 1927a, 1940)) he defended neutral monism. In (1940) he adopted behaviorism explicitly as a methodology, while, as in the earlier works, still accepting introspection as a method. This mix of methods was the basis on which he would turn to the concepts of meaning and truth, drawing in problems of metaphysics. Our focus in this section will be this book, as read through Gödel's eyes.

The picture Gödel assembled was this. The first edition of *Principia* offered a unified correspondence theory of truth, covering both empirical and logico-mathematical knowledge, the MRTJ. Designed to ground the hierarchy of orders, the MRTJ was based upon an account made out in terms of "judgments of perception". After discussions with Wittgenstein, Russell had surrendered the unified approach to truth as correspondence that appeared in *Principia*. Russell's notion of *propositional function*, when taken as a primitive, had led to the collapse of the notions of logical and mathematical truth to the notion of linguistic meaning in Wittgenstein's hands—or so Gödel thought, and not without justification. For Russell took Wittgenstein that way, as did Schlick and members of the Vienna Circle, including Carnap.

<sup>85</sup> Stevens (2005, 2008).

Under the influence of Wittgenstein—as he acknowledged in footnotes—Russell granted as of 1918 that logical (and mathematical) truths were “tautologies” (cf. (1940: p. 140)), i.e., features of symbolism. Yet, as Gödel would have been able to appreciate, Russell’s adoption of Wittgensteinian terminology was always only half-hearted, and never based on any serious philosophical argument.<sup>86</sup>

Russell’s chapter on “Truth and Falsehood” in *The Analysis of Mind* (1921) had been cautious about the sufficiency of the view that mathematics consists of tautologies (II.1):

I do not wish to assert that this is the whole truth about mathematical propositions, for the question is complicated, and I do not know what the whole truth is. But I do wish to suggest that the feeling of self-evidence in mathematical propositions has to do with the fact that they are concerned with the meanings of symbols, not with properties of the world such as external observation might reveal.

By the time he wrote *Analysis of Matter* (1927a), Russell was more confident. In applying the notion of “tautological” to mathematics, he was going even beyond Wittgenstein—just as Carnap and the Vienna Circle had done. So it is striking that in this book Russell made a remark that Gödel *himself* took the trouble to excerpt into his notebook in German on October 7, 1932. Wrote Gödel, working from Grelling’s translation (1927a: ch. XVII),

Philosophy of Matter p. 177 §80: Logical propositions say that two signs are only different ways of presenting the same thing to be expressed [*Logische Sätze besagen, dass 2 Zeichen nur verschiedene Weisen darstellen, dasselbe auszudrücken*].

The context of the quote confirms our picture of Gödel’s understanding of Wittgenstein’s influence on Russell. For Russell had written, in the surrounding passage, the following (1927a: p. 171):

When Kant argued that “ $7+5=12$ ” is synthetic, he was using the subject-predicate definition, as his argument shows. But when we define an analytic proposition as one which can be deduced from logic alone, then “ $7+5=12$ ” is analytic. On the other hand, the proposition that the sum of the angles of a triangle is two right angles is synthetic. We must ask ourselves, therefore: What is the common quality of the propositions which can be deduced from the premises of logic?

The answer to this question given by Wittgenstein in his *Tractatus Logico-Philosophicus* seems to me the right one. Propositions which form part of logic, or can be proved by logic, are all *tautologies*—i.e. they show that certain different sets of symbols are different ways of saying the same thing. Suppose I say: “If  $p$  implies  $q$ , then not- $q$  implies not- $p$ ”. Wittgenstein asserts that “ $p$  implies  $q$ ” and “not- $q$  implies not- $p$ ” are merely different symbols for one proposition: the fact which makes one true (or false) is the same as the fact which makes the other true (or false). Such propositions, therefore, are really concerned with symbols. We can know their truth or falsehood without studying the outside world, because they are only concerned with symbolic manipulations. I should add—though here Wittgenstein might dissent—that all pure mathematics consists of tautologies in the above sense.

In (1940) Russell still defended both the hierarchal treatment of truth—though not Tarski’s, as we shall soon explain—and a correspondence theory of truth for empirical knowledge. As he would later write (1959: p. 98), “Many philosophers speak critically

<sup>86</sup> Dreben and Floyd (1991).

of the ‘correspondence’ theory of truth, but except in logic and mathematics, no other theory had any chance of being right”. His overarching aim in (1940) was to insist, against the Vienna positivists and the pragmatists, that a step into metaphysics is necessary to treat the concepts of truth and meaning.

The theory of meaning and truth Russell developed in (1940) is designed for empirical knowledge, quite appropriately given his audience and targets, and it orbits in the sphere of psychology. Yet, as Gödel surely appreciated, Russell’s arguments included some that broached the foundations of logic and mathematics, i.e., the infinite. Most important of all, Russell remained, despite all the changes in his thought, devoted to *the* fundamental question of philosophy: the nature of truth.

On the one hand, Russell argued that the very notion of logic, “the whole conception of words and sentences as opposed to verbal and sentential utterances”, is “incurably Platonic” (1940: p. 58). On the other hand, Russell remained cautious and circumspect about the foundations of mathematics, praising Gödel’s arithmetization of syntax, but noting the difficulties of accepting the infinite as “given” (p. 71-2). Indeed: such embrace of the infinite is difficult for anyone believing, as Russell said he did, that logical and mathematical truths are “tautologies”.

Even so, Russell returned at the end of the book to truth: the notion of correspondence, the relation of the formal treatment of truth to its nature, to belief and judgment, the distinction between negation and falsehood, judgments of perception, and the law of the excluded middle. These are all discussed and defended in (1940), as of old, the escape from false objectives now worked out in terms of what a sentence “signifies” or “expresses”, a “proposition” being construed as a complex psychological occurrence, and a “fact” as what a sentence “asserts”, i.e., the fact that would make it true.<sup>87</sup>

Gödel (1944) is a meditation on these instabilities within Russell’s philosophy, an effort to make Russell’s philosophy coherent with itself. Gödel’s picture is that Russell’s original *Principia* unified correspondence theory of truth should be revitalized, because “there is no need to give up hope” in the development of a completed philosophy and logic that would allow a defensible step to the infinite as, in some way, “given” (cf. p. 152). This step, if defensible, would circumvent the *Tractarian* collapse, giving philosophers more to say than simply that truth manifests itself in our saying what is the case, *Punkt*. Gödel believed, on the basis of his proposed embrace of an infinitary MRTJ, that the defense could be made out in terms of a theory of (infinitely complex) conceptual complexes, and that this would require, in part, pure philosophy. The notion of “concept” would then analyze the Russellian notion of “propositional function”, while unified notions of “experience” and “truth”—that is, a unified form of the MRTJ—could be retained. The better angels of Russell’s nature, with respect to the nature of truth and falsehood, might then be brought back into philosophy.

There is more than one irony in the fact that Kreisel reported (1983: p. 300) that in conversation with Wittgenstein in the 1940s,

One day after brief, reasonable explanations of the Gödel incompleteness proof Wittgenstein spoke with complete enthusiasm. Gödel must indeed be an extraordinarily original mathematician, for he had derived from utterly banal arithmetical sentences—

<sup>87</sup> Stevens (2008, p. 140-6).

implicitly [Wittgenstein meant] metamathematical sentences—qualities such as consistency. Gödel had discovered a brand new proof method.

## 5.1 Neutral monism

Russell delivered the 1940 William James at Harvard with Carnap, Tarski and Quine in the audience. Gödel understood, as he read the published version (1940), that Russell was attempting to do two things in this venue. First, he aimed to intervene in the then-current analytical philosophical landscape, especially with the logical positivists, Tarski and Quine, but also with the later Wittgenstein, whose ideas were being passed about in *Samzstat*, not yet published.<sup>88</sup> Russell intended to offer a corrective to the anti-metaphysical and nominalistic tendencies he felt too dominant in their work, precisely by focusing on the notions of truth and meaning.<sup>89</sup>

Second, paying homage to the philosopher for whom the lectureship was named, Russell forwarded a treatment of the mental indebted to William James's doctrine of "neutral monism".<sup>90</sup> James had argued (1904) that consciousness does not exist as a quality, property, or directed phenomenon associated intrinsically with mind, independently of its material relations. Instead, experience *contains* relations, rather than intrinsically mental, intentionally-directed sensations that require unification or further projection by the mind. For "mind" and "matter" are simply two different ways of describing what is, so far as its intrinsic properties goes, the underlying "neutral" stuff. Consciousness disappears, on this view, into a complex of relations and descriptions of events, "the name of a nonentity" that "has no right to a place among first principles" (p. 477).

In later discussion of the view, the term "neutral" dropped out,<sup>91</sup> but the aim of an analysis of belief in terms of complexes treated from a monistic metaphysical point of view has of course remained very much alive. The whole issue is bound up with the question of intentional objects and their status. As Gödel may have either been told, or worked out for himself, until 1918 Russell held a view opposed James's, following Brentano and Meinong in holding that the mark of mentality is intentionality, the directedness of experience. On this view, consciousness contains essentially the elements of act and object—though *via* the theory of descriptions Russell had rejected Meinong's notion of the content as an intermediary between

88 Russell framed his comeback to philosophy by publishing (1936), which ends criticizing Wittgenstein's finitism. Applying his arguments to Wittgenstein's student Alice Ambrose, who had—against Wittgenstein's wishes [Monk (1990: p. 214); cf. McGuinness (2008: letters 187-188)]—claimed to publish views indebted to Wittgenstein, Russell held that it is not logically impossible to run through the whole expansion of  $\pi$ , but "*medically impossible*" (p. 143). It is not impossible that Gödel had read this essay at the time he wrote (1944). On Carnap's contributions to Russell's manuscript of (1940) during discussions at the University of Chicago in 1938-1939 see Monk (2000: p. 220ff).

89 In (1940: p. 82) Russell mentioned the treatment of color-incompatibility in terms of the notion of a "grammatical" truth. While he did not wish deny this account, he said, he was not sure what it meant. One may compare Russell (1936: p. 140) where the same point is made in connection with both Wittgenstein and Carnap.

90 The term "neutral" embracing "monism" was invented by Sheffer, in his (1908). For the term "neutral" as applied to what is scientific, see Russell (1917: ch. 1). Russell attempted to add "neutral" forms to his (1913) under the pressure of Wittgenstein's questions.

91 Cf. Stubenberg (2013).

mind and object. During his shift into epistemology after *Principia* Russell argued against James's neutral monism (1913). But afterwards—that is, after the influence of Wittgenstein and the American behaviorists on his views about consciousness—he came to accept James's view.<sup>92</sup>

Many philosophers reading Russell's (1940) have dismissed it as the work of an aging and uninteresting philosopher, a book that engendered, through its influence on Quine, a form of psychologistic reductionism and emotivism that damaged the analytic tradition.<sup>93</sup> In Quine's own account (1985: p. 150), the intellectual stimulation of Harvard in 1940 came from discussions with Tarski and Carnap, and not with Russell.

Yet the influence of Russell's lectures on Quine is clear (cf. Frost Arnold (2013)). Russell's (1940) offers a vivid illustration of how, from a certain set of “common factor” premises, it follows ineluctably that two observers cannot directly perceive the same object.<sup>94</sup> The future development of Wittgenstein's, Carnap's, Tarski's and Quine's work was to shape the next generation of philosophers, and Russell (1940) is an ideal textbook for working out why they would have developed different views. In particular, it vividly illustrates the attractiveness of accepting as fundamental (as Tarski and Quine did) the Tarskian approach to truth as primarily, rather than derivatively, a property of sentences and/or their utterances, and not of beliefs. It also perfectly motivates, in a negative way, many of the subsequent moves that Quine, Wittgenstein, Sellars, Davidson and Putnam would make against, e.g., solitary sensation language—which is advocated by Russell explicitly for singular judgments (p. 186).

By contrast, while writing (1944) in 1942-1943, Gödel was happy to see Russell attempting to put truth—analyzed in terms of beliefs, or believed-sentences, rather than sentences *simpliciter*—back into the center of discussion. And he welcomed Russell's (1940) criticisms of verificationism and nominalism, as well as his speculative approach to the mind. Like Russell, Gödel regarded the increasingly popular philosophical approach to truth based upon Tarski as a trend that missed certain basic philosophical questions. Gödel's (1944) sought to draw Russell's whole approach to truth back into the center of philosophy by reframing his arguments about correspondence in terms of the older *Principia* account of “judgments of perception”, but now insisting on the infinitely complex character of conceptual experience.

In (1940) Russell made explicit his differences with logical positivism. He praised the positivists' contributions to knowledge and their cooperative, scientific spirit—after all, he had himself engendered this very way of doing philosophy. Yet he argued that their forms of nominalism and empiricism were inadequate. Rehearsing the history of his moves toward constructivism under the influence of Wittgenstein and Ramsey in the second edition of *Principia*, he forwarded an alternative psychologistic and naturalistic basis for empirical knowledge. Here, Gödel feared (1944: p. 130), Russell was perhaps turning toward a “merely psychological investigation of the actual processes of thought” which would be of little help in assessing the foundations of logic.

However, even in these lectures Russell pursued logical, as well as psychological investigations. Still arguing for the elimination of “negative facts”, he accepted

92 Cf. Russell (1921), (1959: ch. XII).

93 Wang (1986), Monk (2000: ch. 8).

94 Putnam (2013). More sympathetic accounts of Russell may be found in Levine (2009), Stevens (2005).

only positive atomic facts, distinguishing between what a statement “signifies” or “expresses” (a proposition, here construed as a complex psychological occurrence involving images, reactions, and so on) and what it “asserts” (the fact that would make it true) (cf. Gödel (1944: p. 129, n. 7)).<sup>95</sup> Most interestingly from Gödel’s point of view, Russell departed both from James’s pragmatism and from Carnap’s conventionalism in his analysis of truth. In the closing chapters of (1940), Russell went so far as to insist that a step into metaphysics is necessary to resolve what was always for him “*the fundamental question of philosophy*”, the nature of truth.

In reading Gödel, then, one must not impose Quinean or Tarskian motifs upon his remarks.<sup>96</sup> And one must never forget the continuities in Russell’s thought to which Gödel is responding. Gödel’s (1944) was written to urge Russell back into the fold, embracing the better angels of his nature, a proper and direct concern with metaphysics and a *unified* treatment of truth as correspondence contained within a conceptual complex. Having already worked on transfinite orders as indices for the theory of types, Gödel sought to find a way to return—by generalizing—to *Principia*’s MRTJ.

Overall, Russell argues two main points in (1940). First, the linguistic articulation and conceptualization of experience, as analysis, always leaves a conceptualizable residue behind—and hence cannot be viewed as constituting an ultimate basis for knowledge to which an essentially pragmatic or purely conventionalist element (such as James’s or Carnap’s) can be added to analyze truth. The argument turns, interestingly enough, on the (general) undefinability of truth, and on the distinction between “correctness”—which is, Russell says (p. 186), a “social” concept—and truth, which is not. Second, the possibility of our knowledge of truth requires a step into metaphysics beyond pure empiricism, an embrace of (complex) universals, because the analysis of truth entails a demand for an analysis of the notions of belief and statement that allows for truth as correspondence.

As described in §1, Gödel proved the undefinability of truth in 1930, as Tarski acknowledged. In (1940) Russell did not take Tarski to have elucidated the fundamentals of truth. He noted (p. 62) that he himself had anticipated Tarski’s hierarchy of languages, *via* an informal argument in the Introduction to Wittgenstein’s *Tractatus*. And then Russell went on to develop his own updated version of the hierarchy of languages, insisting that it was “not identical with Carnap’s or Tarski’s”, nor was it “the only possible one” (p. 63n). His motivation, as Russell nicely encapsulated the point in a retrospective review of his own philosophical development (1959: p. 136), lay in the MRTJ approach:

Truth and falsehood both belong primarily to beliefs and only derivatively to sentences. Beliefs, if they are sufficiently simple, can exist without words, and there is every reason to suppose that they exist in the higher animals. A belief is “true” when it has an appropriate relation to one or more facts, and is false when it does not have such a relation. The problem of defining “truth”, therefore, consists of two parts: first,

<sup>95</sup> Cf. Stevens (2005, p. 140).

<sup>96</sup> As Parsons suggests (1990: p. 112), a Quinean distinction between ontology and ideology, resting upon the Tarskian apparatus, is indeed foreign to Gödel, just as it was to Russell. Note that Parsons is careful, in raising this question, to insist that propositional attitude expressions are not included in the proposed Tarskian account. They certainly are, however, both in Russell’s work after (1921), and in Gödel’s 1942-3 *Max Phil IX-X* notebooks.

the analysis of what is meant by “belief”; and then, the investigation of the relation between belief and fact which makes the belief true.

In (1940) Russell reasoned that (p. 63) “the hierarchy must extend upwards indefinitely, but not downwards, since, if it did, language could never get started. There must, therefore, be a language of lowest type”. In fact, according to Russell, “at the lowest level of speech, the distinction between sentences and single words does not exist” (p. 29).<sup>97</sup> In describing this lowest level “ $n$ th level language”, which he calls “the object-language” or the “primitive language”, Russell insisted—unsurprisingly for any attentive reader of *Principia*—that in it neither denial nor negation nor the predicates “is true” and “is false” can occur (p. 63):

It is clear, from Tarski’s argument, that the words “true” and “false” cannot occur in the primary language; for these words, as applied to sentences in the  $n$ th language, belong to the  $(n+1)$ <sup>st</sup> language. This does not mean that sentences in the primary language are neither true nor false, but that, if “ $p$ ” is a sentence in this language, the two sentences “ $p$  is true” and “ $p$  is false” belong to the secondary language. This is, indeed, obvious apart from Tarski’s argument. For, if there is a primary language, its words must not be such as to presuppose the existence of a language. Now “true” and “false” are words applicable to sentences, and thus presuppose the existence of language. . . In the primary language, therefore, though we can make assertions, we cannot say that our own assertions or those of others are either true or false.

. . . When I say that we make assertions in the primary language, I must guard against a misunderstanding, for the word “assertion” is ambiguous. It is used, sometimes, as the antithesis of denial, and in this sense it cannot occur in the primary language. Denial presupposes a form of words, and proceeds to state that this form of words is false. The word “not” is only significant when attached to a sentence, and therefore presupposes language. Consequently, if “ $p$ ” is a sentence of the primary language, “not- $p$ ” is a sentence of the secondary language.

This recalls the general approach to  $n$ th truth in *Principia*. Given that he was drawing psychology into the discussion, however, Russell was quick to clarify (p. 63):

It is easy to fall into confusion, since “ $p$ ”, without verbal alteration, may express a sentence only possible in the secondary language. Suppose, for example, you have taken salt by mistake instead of sugar, and you exclaim “this is *not* sugar”. This is a denial, and belongs to the secondary language. You now use a different sprinkler, and say with relief “this *is* sugar”. Psychologically, you are answering affirmatively the question “is this sugar?” You are in fact saying, as unpedantically as you can: “the sentence ‘this is sugar’ is true”. Therefore what you mean is something which cannot be said in the primary language, although the same form of words can express a sentence in the primary language.

Hence Russell eliminated from his “object language” all logical connectives, all comparatives, and all propositional attitude expressions (p. 64f), for they already

<sup>97</sup> It is hardly an accident that the opening language games of *Philosophical Investigations* (2009: §§1-2) contain no clear distinction between sentences and single words and no explicit form of negation or truth predicate. Wittgenstein’s slow, step by step, approach the notions of negation and assertion through variations of possible scenarios is intended to answer Russell’s psychologism with a new approach to logical method. Quine too, in (1992) drew a distinction between taking a sentence “holophrastically” (not divvied up by analytical hypotheses into words) and taking it “analytically” (after analysis), in a book whose title is drawn from Russell (1921: ch. XIII).

presuppose language. And he defined “object-words” from both a logical and a psychological point of view. Logically, they are those words taken as having (p. 65)

. . . meaning in isolation, and, psychologically, as words which have been learnt without its being necessary to have previously learnt any other words. These two definitions are not strictly equivalent, and where they conflict the logical definition is to be preferred. They would become equivalent if we were allowed to suppose an indefinite extension of our perceptive faculties. We could not, in fact, recognize a chiliagon by merely looking at it, but we can easily imagine beings capable of this feat.

Given his results with transfinite ordinal indexing, Gödel would have been receptive to the idea of an “indefinite extension of our perceptive faculties”. But once Russell turned to a class of “possible” object-words, Gödel would also have been inclined to think that at the primary point certain mathematical and/or logical concepts would need to be presupposed in our very understanding of the object-language. Russell continued (p. 65):

On the other hand, it is clearly impossible that any being’s knowledge of language should begin with an understanding of the word “or”, although the meaning of this word is not learnt from a formal definition. Thus in addition to the class of actual object-words, there is a class of possible object-words. For many purposes the class of actual and possible object-words is more important than the class of actual object-words.

The following remark of Russell’s expresses the idea that there can be understandings that cross over levels of the hierarchy of truth (p. 69):

What kind of simplicity makes the understanding of a word into an example of understanding an object-language? For it is to be observed that a sentence may be spoken in the object-language and understood in a language of higher order, or vice versa. If you excite a dog by saying “rats”! when there are no rats, your speech belongs to a language of higher order, since it is not caused by rats, but the dog’s understanding of it belongs to the object-language. . . Whenever you doubt or reject what you are told, your hearing does not belong to the object-language; for in such a case you are lingering on the words, whereas in the object-language the words are transparent, i.e. their effects upon your behavior depend only upon what they mean, and are, up to a point, identical with the effects that would result from the sensible presence of what they designate.

In *Max Phil* IX [17-18] Gödel criticizes Russell’s hierarchical treatment of understanding by emphasizing that even at the primary level it is essential to take into account the logical role of a proposition, thereby taking it in as a conceptual complex. An atomic proposition—whether or not it is of the form of a propositional attitude ascription—must be treated in terms of the logical role it plays in structuring of the totality of possible thoughts. Writes Gödel,

*Remark (Philosophy):* According to *Russell* the signification [*Bedeutung*] of a proposition in assertions and as argument of truth functions is the same and is different from its signification in “A believes, wishes, claims, *etc.* p”. Namely, in the first case [*transparent occurrence*]<sup>98</sup> something is said through it, in the second case about it. In the first case the signification is no unity, but the plurality of constituents combine through what is asserted. In the second case it is the series of symbols [or more precisely its “form”]. In

<sup>98</sup> This is written in English in *Max Phil* IX.

the first case is it somehow supposed to be the result of the proposition's being-spoken? But in truth the symbol is often believed and claimed and the distinction is not so great.<sup>99</sup> What is correct is that for many *occurrences* to be understood it is necessary that they are not grasped as unities [which can be replaced by an abbreviation], but as a unity of bound pluralities [necessary for the main combining sign of a sentence]. Sometimes it is still possible, sometimes already *in principle* impossible, as e.g. in the proposition (p)[A asserts  $p \supset p$ ].

This objection, as Gödel would have grasped, is reminiscent of the young Wittgenstein's *Tractatus* reworking of the MRTJ: a proposition signifies differently from a name, even at the "primary" atomic level, as it figures in a larger logical space that may be formally articulated. Wittgenstein had objected to Russell's treatment of the MRTJ that an analysis of judgment "must show that it is impossible for a judgment to be a piece of nonsense" (5.5422). Gödel is wondering how Russell's treatment of truth might be construed so as to meet that requirement. Wittgenstein had gestured at an analysis for the MRTJ when he wrote that (5.542) "A believes that  $p$ ' and 'A says  $p$ ' are of the form ' $p$  says  $p$ '—meaning by " $p$ " a reference to some fact in the world (an event of utterance, for example) serving the role of projector—and that "this does not involve a correlation of a fact with an object, but rather the correlation of facts by means of the correlation of their objects". The distinctively Gödelian twist in *Max Phil* IX [17-18] is that there is no sharp distinction between believing (i.e., projecting) a symbol and believing a conceptual complex (to exist). Such is reflected in Gödel's own mathematical practice, as we have seen.

Gödel's numerous discussions of awareness and truth in *Max Phil* IX-X thus indicate familiarity, not only with Russell's form of "neutral monism" and the MRTJ, but also the arguments given in (Russell 1940). The fact is that Russell's thinking about the nature of truth in the period before, during, and after writing *Principia* was thoroughly entangled with his treatment, not only of false belief and negation, but also of perception and the question whether a subject or perceiving self exists as a purely "mental" entity (cf. Russell (1959: p. 139)). And Gödel knew this.

He also knew from at least 1932 that Russell's MRTJ version of the correspondence theory of truth was very different from Schlick's (cf. §6.1). Gödel was aware that Schlick's endorsement of the correspondence theory had earned the disapprobation of some in the Vienna Circle.<sup>100</sup> As a matter of fact, in I §10 of his "Allgemeine Erkenntnislehre" (1925), entitled "What is Truth?", Schlick referred explicitly to chapter XII of Russell (1912), where a version of the MRTJ is defended that is somewhat idiosyncratic in being based on the notion of an "act" of judging. Schlick argues (p. 68) that although Russell's theory is "ingenious" in being a theory of truth

<sup>99</sup> *Max Phil* X [21f] develops this idea:

*Remark (Grammar):* The identification of the sign with the signified [note: signs in the broadest sense, where concepts and representations are also signs] is a characteristic feature of natural language, e.g.,  $\in$  with  $=$ , A is true  $\equiv$  A is actual, in the sentence " $(\lambda x) \varphi(x)$  doesn't exist" means  $(\lambda x) \varphi(x)$  on the one hand the symbol (or the concept) and insofar one can speak of it, on the other hand that which is indicated through the concept and insofar it doesn't exist. [note: this perhaps explains the *Platonic Paradoxes of the false*. [We only perceive pictures to be true, and nothing actual.]

<sup>100</sup> Oberdan (2013: §8). Stadler (2001) stresses the diversity of views on the notion of truth within the Vienna Circle.

“built up entirely out of differences that characterize various kinds of relations”, his own theory is superior, because it “rests solely on the relation of pure coordination or correspondence, which is the simplest and most general of all relations”.

Both Schlick and Russell explicitly rejected the coherence, pragmatist, and phenomenological accounts of the nature of truth. Both accepted the use of the notion of “concept” in articulating their theory. And both rejected “false objectives”, intensional (and intentional) entities that correspond to false judgments. In Schlick’s “On the Nature of Truth in Modern Logic” (1910) an original view of truth as univocal designation had been developed, on which a judgment, as a structured complex of its constituents, is taken to be coordinated with the fact consisting of the entities signified by the judgment’s constituents. Later, in his (1925) Schlick argued that concepts are formed in clusters, just as the concepts of a mathematical field are defined in the context of an axiomatic theory.<sup>101</sup> Schlick’s theory takes general laws to be primary, rather than particular judgments. For on his view, terms should be interpreted in the context of a theory so as to designate objects and properties, and then sentences would designate (or fail to designate) facts. Yet this view lends itself—as Russell’s, beginning with particular “judgments of perception”, does not—to fairly straightforward adaptation to a conventionalist point of view and the elimination of the notion of “correspondence”. This Gödel knew from firsthand experience, having sat in on the Vienna Circle’s discussions of truth.

Ever since his salad days in Vienna (cf. §6.1 below), Gödel had understood that to be persuasive, the correspondence theory of truth would have to be complex, and not simple and general, the way Schlick’s had been. Thus Gödel did not simply follow Russell in accepting a general analogy between grasp of fundamental concepts and sense-perception. Instead, he took Russell to have bequeathed the right approach to truth and knowledge, one awaiting resolution by philosophical and logical means.

Despite his hesitations about Russell’s treatment of meaning, Gödel surely followed him in accepting the need for a hierarchical treatment of truth. Moreover, he would have applauded Russell’s insistence, against the positivists, that there is a fundamental distinction between “truth” [*Wahrheit*] and “correctness” [*Richtigkeit*] (p. 186):

It would seem. . . that “correct” cannot be used in defining “true”, since “correct” is a social concept, but “true” is not.

This distinction grounded, in fact, Russell’s rationale for positing “solitary language” (p. 186):

I think that, in fundamental discussions of language, its social aspect should be ignored, and a man should always be supposed to be speaking to himself—or, what comes to the same thing, to a man whose language is precisely identical with his own. This eliminates the concept of “correctness”. What remains—if a man is to be able to interpret notes written by himself on previous occasions—is constancy in his own use of words: we must suppose that he uses the same language today as he used yesterday. In fact, the whole residuum of what was to have been done by the concept of “correctness” is this: speaker and hearer (or writer and reader) must use the same language, i.e. have the same interpretive habits.

---

<sup>101</sup> Oberdan (2013).

*Max Phil IX-X* are laced with a variety of investigations of the difference between “correctness” and “truth” for this reason. However, it is striking that unlike Wittgenstein, Gödel nowhere attacks the roots of Russell’s methodological solipsism. Instead, he appears to have embraced that methodology, and that embrace appears to have remained with him for a long time afterwards, even after his turn toward Leibniz and phenomenology. As we read in notes taken down by Sue Toledo in conversation in 1972 (2011: pp. 200-201), Gödel reportedly said that

The understanding of the system of primitive terms and their relationships cannot be transferred from one person to another. . .

There are probably different ways to find these primitive terms. The method of phenomenology is not to investigate the terms themselves, but rather to investigate how we handle them. (Husserl never mentions that his goal for phenomenology is finally to come to an understanding of the primitive terms themselves.)

Another thing he doesn’t mention is that when we investigate our way of handling these concepts, there is no reason to assume that we always handle them correctly. If one considers the early age at which we start, it would not be surprising that [we] would make many wrong combinations of the primitive terms. . . Husserl’s *epoché* on the other hand is essentially an exclusion of criticism, of any concern about truth and falsehood.

As Gödel appreciated, it was Russell—and not only Husserl—who again and again insisted that at the foundations truth must attach directly to belief and actual experience. On this approach particular experience, and not language, is primary, although language, and the notion of a statement, must figure centrally in the logical theory at some point. As Gödel saw it, Russell rightly always insisted on a correspondence theory of truth, as Husserl had not. This is because Russell faced head on, at least in his earlier writings, the question about non-demonstrative knowledge which occurs in connection with the indefinables (e.g., the notion of truth) and argued for a correspondence theory for it.

What could be more appropriate for fleshing out the analogy between our grasp of the indefinables and sense-perception? Perception, whether sense-perception or “perception” of a fundamental concept, does not come packaged with sentential or propositional labels. And it is always unique, however hazy the contours of what we may truly say of it may be. Gödel always agreed with Russell that the analysis of truth, depending upon that *Ur*-point, must be capable of yielding recognition of *particular* concepts, or at least of particular conceptual structures, especially when it comes to the indefinables. It falls to logic, to philosophy, to make sense of the very idea of judging, truly or falsely. If analysis is to shape our understanding of truths, and be possibly correct (or incorrect), and we are to think in terms of a hierarchy for truth, then, at least with respect to the indefinable notions, we must, it seems, just “see” them in action in our statements, rather than deduce them.

The work Russell had to do in his (1940) was to make sure that his naturalistic approach to experience and the making of statements in language could deliver something to play the role of his earlier notion of “acquaintance”, for this had encapsulated the idea of perception coincident with truth, an epistemic, immediate, and direct access to constituents involved in understanding, concepts, and logical form. Offering a sweeping overview of then-current psychological theories, from Gestalt theory to behaviorism, Russell developed a very complicated amalgam of notions in (1940).

Yet the danger Russell's approach had to face—unlike Husserl's—was that the role for such particular experiences would vanish altogether from the foundations of logic. This danger is precisely what Wittgenstein had, after all, urged on Russell in 1912-1913. In the *Tractatus*, written in the wake of his conversations with Russell, Wittgenstein laid down his conception of a truth condition as a kind of challenge to Russell to do better than “showing” for “the simplest kind of *truth* and *falsehood*”. Logical forms, as the *Tractatus* has it, are shown, not said, i.e., they are not constituent “bones” of propositions, but a kind of scaffolding by means of which what we say is erected.

A sign of Gödel's appreciation of the philosophical problem space—perhaps the most important to be verified in *Max Phil* IX-X *vis-à-vis* Russell—is that he is aware that in broaching concept theory at the end of (1944), he would have to come to terms with the difficulties Russell had earlier faced, in 1913, in discussions with Wittgenstein about the MRTJ. Although Gödel could not have known of the detailed inner history of Russell's abandonment of his 1913 manuscript *The Theory of Knowledge* under the pressure of Wittgenstein's objections, Gödel certainly knew *Principia*, the *Tractatus*, and enough of Russell's subsequent writings well enough to have grasped the difficulties.

Russell's repeated mention of Wittgenstein certainly did concern the issue of constructivism vs. Platonism in the foundations of mathematics. But Gödel knew that there was something deeper going on, philosophically speaking. Gödel believed that Wittgenstein had deeply influenced Russell's approach to truth, leading him to abandon the unified approach adopted in the first edition of *Principia*, and embrace a constructivistic theory of orders in the second edition.<sup>102</sup> Gödel's invocation of Leibniz in the closing coda to (1944) was intended to complement and to challenge Russell, evoking a response (cf. §6.2). Without that response, Gödel feared that the purely philosophical aspects of his essay would remain incomprehensible. And so, standing in isolation, they are.

## 5.2 The inifinitary MRTJ

Gödel's construal of the MRTJ, unlike Russell's, construed the constituents of the conceptual complex existing in the case of a true judgment as *projectively* and *internally* related, both to one another and to whatever complex or constituent would play the role of the judging subject. Infinite internal conceptual complexity of elements, the whole conceptual complex of true judgment-facts, would be taken “as given” from the start, and an unbounded articulation of “inner” correspondence (in mathematical terms, relative consistency results) would emerge. Truth would then be construed,

<sup>102</sup> Landini (2007) argues that Russell made the move to the second edition more or less on his own. Other views have been proposed, such as Levine (2013). It is not our task here to adjudicate the vexed question of to what extent Wittgenstein was Russell's apprentice, or *vice-versa*, and in what ways. Useful pieces relevant to this important controversy include: Pears (1977), Griffin (1985), Hylton (1990), Ricketts (1996), Carey (2007), Pincock (2008), Potter (2009), Zalabardo (2012). It is uncontroversial, whatever the account of the details, that Wittgenstein did influence Russell to adopt the terminology of “tautology” (Dreben and Floyd (1991)).

as in *Principia*, in terms of correspondence, but in a “reflexively reflected” way.<sup>103</sup> Judgment and belief facts would be unfolded through infinitary ordinal indexing, in the case of mathematical truth; in more general terms, there would be no bound to the complexity of our capacity for true judgment. Yet also, Gödel noted in the margin of *Max Phil IX* [21], “There are no negative facts, because that would be a limit”.

On this view, mathematical truth manifests itself in our “analysis” of “experience”, but truth is not eliminated as a redundant property, even if, in an extended sense, mathematical truth may be said to be “analytic” (where the latter notion is, in general, undefinable in the sense of being “irreducible to anything more fundamental” (Gödel (1944: p. 151)). Instead, the correspondences are regarded as conceptual complexes already there to be judged, but requiring “projection” of some kind for our knowledge of them. As to epistemology, this judging would be everywhere controlled by proof and the axiomatic method.

On this approach, however it is developed, truth has to bottom out in particular “judgments of perception”, i.e., simple elements of particular judgments (conceptual complexes). Gödel, like Russell, is aware of the epistemological difficulties at the base, and regards them as properly philosophical, and not mathematical. *Max Phil IX* [15] remarks explicitly that with regard to the most fundamental concepts of religion and philosophy, mathematics is “misleading”.

As to pure philosophy, Gödel quite properly adduced Plato’s *Theaetetus*, with its fundamental questions about knowledge, perception, false objectives and simples—just as would Wittgenstein, unbeknownst to Gödel, in *Philosophical Investigations* (§46).<sup>104</sup> Equally relevantly, Russell himself mentioned the dialogue at the beginning of chapter VIII of (1940), “Perception and Knowledge”, as follows:

The word “perception” is one which philosophers, at an early stage, took over, somewhat uncritically, from common sense. *Theaetetus*, when Socrates asks him for a definition of “knowledge”, suggests that knowledge is perception. Socrates persuades him to abandon this definition, mainly on the ground that percepts are transient, whereas true knowledge must be of something eternal; but he does not question the occurrence of perception conceived as a relation between subject and object. To common sense it seems obvious that we perceive “things”, at any rate with the senses of sight and touch. Sight may, on occasion, be misleading, as in the case of Macbeth’s dagger, but touch never.

Russell’s attack on the dual picture of belief—as well as the idea that we can “perceive” what is not there—derives, as we have seen, from *Principia*, and in this chapter of (1940) he gives familiar arguments, now buttressed by psychology, to show, as he wrote (p. 188), that “if we see sun-spots, there are sun-spots”. Gödel would investigate at length the issues of lies and dreams, commenting on Russell’s discussion in his *Max Phil* notebooks: Russell argued in (1940: ch. XIV) that lies cannot be part of the primary language, for lies involve reflection, and “when language is spontaneous it cannot lie”.

<sup>103</sup> Putnam (1985) invokes Gödel’s incompleteness theorem to forward an epistemic argument against reductive computationalism. Though the issue is construed in terms of knowledge—in this context, justification in a broad sense—truth is part and parcel of the backdrop to Putnam’s argument.

<sup>104</sup> Wittgenstein’s remark was drafted by 1936, in the first version of *Philosophical Investigations*.

This whole discussion is of course intimately connected with the infinitary MRTJ, for the issue is how it might be that by “noticing” particular conceptual complexes, we can, in one sense or another (to use a phrase critically examined in Plato’s *Theaetetus*), “run through the elements of the whole”. In (1944) Gödel is, after all, seeking a conceptual basis for accounting for the paradoxes. At *Max Phil IX* [34-5], after discussing Russell’s analysis of time-consciousness, Gödel wrote:

*Remark (Philosophy): Theaetet[us]:* (1.) A thing can appear to be, without being [and even so is asserted]. How is that possible? (2.) One not only designates [*benennt*] with a name, but also does something as one binds names with verbs and [35] names this “thought” (*discourse*).

What is characteristic of Gödel is that he immediately connects this problem, as had Russell in the *Principia*, to the definition of truth. At *Max Phil IX* [35] we read:

*Remark (Philosophy):* The definition of truth is correct [*richtig*] in the *transcendent* theory, but it must give an immanent truth-criterion and to this counterpart the *Transcendent Definition* must appear as a “useless tautology” [similar to Carnap’s definition of “analytic”].

Gödel is aware that in Tarski’s analysis of truth, T-sentences must appear to be “useless tautologies” from below, and here he adduces Carnap’s phraseology in *The Logical Syntax of Language* (1934: §78f). Carnap’s position derived from the *Tractatus*, as well as his understanding of Gödel’s arithmetization of syntax. For Wittgenstein and Ramsey—who followed Wittgenstein on the point (1923, 1927)—to say “*p* is true” at the simplest level adds nothing beyond saying “*p*”, since the analysis of “I believe *p*” is just “*p* says *p*”, that is—to say what is said—“*p*”. The philosophically least misleading thing to say is that the locution “*p* is true” should be banished insofar it express an apparent move to a transcendent or fact-based *explanation* of the truth is taken to of *p*. Instead, the best way of isolating the specific content of *p* is to just say “*p*”, thereby drawing truth (and falsehood) in, as Wittgenstein had written in the *Tractatus* (5.541).

On this view, the move to molecular and quantificational structures and to higher level orders of truth is acceptable so long as this is regarded as essentially a series of parametrizations, index steps to avoid contradiction, something purely formal and linguistic, and hence, an activity of arrangement not really *adding* to the content of what may be said.

Now in the *Principia* Russell too is very clear that his types and orders must be finite; he regards them, ultimately, as a kind of linguistic device.<sup>105</sup> As Gödel remarks, one must “give an immanent truth-criterion”, and from this point of view the “Transcendent Definition” must appear as a “useless”. Of course Gödel rejects this appearance; for him the theory of types and orders is substantive, not something merely linguistic. For Gödel, who had already devised applications of *L*, the transcendent move to a higher type level is intrinsically impredicative and ontological, though liable to *appear* as a merely tautological move from below. The position we are attributing to Gödel is subtle, and not dogmatic. For truth as correspondence only falls into place within proof, the context of something purely mathematical.

<sup>105</sup> Cf. Landini (1998, 2007).

As we have said, Gödel was perfectly aware that the *Tractatus* intended to resolve at least some of the difficulties with the *Principia* MRTJ by construing form as the *possibility* of structure, thereby denying that it plays the role of any actual object or constituent. Gödel did not, however, think it necessary to follow Wittgenstein hook, line and sinker in drawing out the consequences of this point. With the help of Leibniz, another result might be reached. In *Max Phil* IX [11] the fifth thesis about concepts Gödel wrote down is that “The form of a sentence cannot be a constituent of the sentence, for otherwise it would fall under itself”. His idea, we believe, was to agree with the *Tractarian* conclusion, but to use the limitations of definability with respect to truth to defend an infinitary version of the MRTJ—hence a correspondence theory of truth—for logic and mathematics.

Like the original *Principia* theory, Gödel’s MRTJ would be based on an analysis of particular judgments, taken as giving “the simplest kind of *truth* and *falsehood*” and construed as constituted by analysis of dual perceptual experiences. The generalized MRTJ, a form of concept theory, was to overcome the collapse of mathematical and logical truth into a species of conventional linguistic meaning à la Carnap and make sense of truth as correspondence for the transfinite realm.<sup>106</sup> A task of philosophy, as Gödel saw it in 1942-1943, is thus to transform the *analogy* between knowledge of concepts and sense-perception into a *correspondence*, just as Russell had tried to do in *Principia*’s first edition in offering a definition of “the simplest kind of *truth* and *falsehood*” in terms of correspondence in judgments of perception.

Gödel knew that the self and “awareness” had to be at issue in any such analysis. Although consciousness is not discussed in (1940)—given the embrace of neutral monism, it would not be—the history of Russell’s struggles with *first* truth and negation is rehearsed, and Russell argues explicitly, against the Gestalt theorists, that “there is no theoretical limit” either to the complexity of what can be perceived or to our ability to devise specifications of precisely *what* we perceive in words (1940: p. 153, 355). While Gödel does express periodic hesitations about neutral monism’s treatment of the self in *Max Phil* IX, especially when he explores the role of the first person pronoun, he does see Russell’s approach as the only viable one to truth.

All this is folded into Gödel’s broaching, in his philosophical notebooks, a general *Urprinzip*, a “principle of principles” or original structure of perceivable concepts. This plays the role for Gödel of a kind of analog of the unity of Reason, as Kant would have seen it. But for Gödel the *Urprinzip* constitutes a criterion, not merely of empirical truth, in its systematicity, but of truth as such, in its conceptual unity. The *Urprinzip* allows, in other words, for the recognition [*anerkennen*] of truth, conceived as unfolding in particular things that are said, yet unified as a structure of concepts.

For Gödel, as for Frege and the early Russell, there are no modes of truth, there is only one notion of truth. A truth condition is neither a property or state of a thought or sentence, nor a conditional statement involving articulation of a possible fact. It is, literally, an existing conceptual structure. The stratification of our linguistic articulation of truth, what the *Principia* called the “systematic ambiguity” of truth in terms of types and orders, pertains to our knowledge, but not to truth as such. Crucially,

<sup>106</sup> In Gödel (1944) the word “tautology” does not occur, only “analytic” does, although of course not in Wittgenstein’s or Carnap’s sense, but in something more like Leibniz’s sense.

“truth as such” cannot be built up constructively, from below, by means of linguistic articulation. It must be presumed in full, right at the outset.

[Struck out on the margin]: Reality possibility. Truth is a relation? Structure of objective Concepts. [On margin but not struck out]: How is it that unique mathematical *problems* are solvable? Assumption of the existence of a systematic procedure *nota notae*.

*Remark (Philosophy)*: There is probably some concept that is just as precise as (or even more precise than) mathematical concepts, and which contains the principle of every structuring of concepts [*Begriffsbildung*] [namely for all the concepts which originate not through abstraction from sense perceptions]. That is, from it alone (if one once had grasped it) systematically all concepts of the series after it would to some extent result “from itself”, together with the relations holding between them [that would be the realizable *constitution* system [*Konstituti<utions->*system]. This concept (the original principle [*Urprinzip*], (note: Is there a word for this in everyday language? Is it a relation concept?) would also have to have the following properties:

1. To be precise.
2. It must bring unity and order into present logical fundamental concepts [*Grundbegriffe*].
3. It must contain a method, a method of tackling a problem, that is, the method provides for a goal (or describes the correct *intelligent* behavior). That is, if one approaches another *problem* with this concept, a “light” must dawn over it.
4. Perhaps this concept is no unity, but a structure of concepts, consisting of several concepts standing in connection with one another.<sup>107</sup>

There is a remarkable resonance of Gödel’s remarks about “light drawing” with some of Wittgenstein’s investigations of aspect-dawning—something perhaps not too surprising when we reflect that both writers were preoccupied with responding to Russell’s conception of truth.<sup>108</sup> Gödel’s introduction of the idea of an *Urprinzip* is, however, distinctive, and he immediately applied it to his rethinking of Russell’s analogy between zoology and logic:

**Remark (Philosophy)**: The highest categories (*Substance*, etc.) are class concepts and are to be distinguished from things, which we seek to control through them (naturally they themselves belong in a category of these things). Perhaps however the classification is not at all the most effective method for controlling the multiplicity of things. [note. Perhaps it is the final distinction between *systematic Zoology* and the explanations of modern science.]

At stake, briefly put, is the unity and singularity of the notion of truth as such, given through the the final correct theory of concepts. While there is a wider history surrounding the idea of a ground of all grounds, a principle of all principles or *Urprinzip*, and the idea of “light dawning” over an organic whole—e.g. in Leibniz,

<sup>107</sup> *Max Phil* IX, [63-64]; cf. [45] for the connection with negation.

<sup>108</sup> Wittgenstein wrote (1971, §139-142):

141. When we first begin to *believe* anything, what we believe is not a single proposition, it is a whole system of propositions. (Light dawns gradually over the whole).

141. It is not single axioms that strike me as obvious, it is a system in which consequences and premises give one another *mutual* support.

Goethe, and Husserl<sup>109</sup>—we see here the demand for the principle arising out of Gödel's reflections on Russell, truth, and the nature of logic.

A fundamental difficulty with Gödel's *Urprinzip*—i.e., the totalization of all possible judgment-complexes and corresponding complexes through all limits of all possible indexings of truth—lies at a point that is structurally analogous, conceptually and philosophically speaking, to the series of difficulties that Wittgenstein had found with Russell's notion of “logical form” in the context of the 1913 (finitary) MRTJ. The objection is as follows. If Gödel's imagined *Urprinzip* is itself to be considered true, rather than false, it will already have to have a negation. If this is so, then the truth for that will have to be clear. But if Gödel's *Urprinzip* is false, then there is a question whether this still counts as a principle that happens to be false—or whether it would constitute chaotic nonsense.<sup>110</sup> Gödel outlines the difficulty at *Max Phil IX* [45]:

*Remark (Philosophy):* If one regards the objects of mathematics as constructed through the mind [*Geist*], one brings in necessarily a temporal element. They [i.e., the objects] then exist first after the construction and one can say with right, with Poincaré, that as to their classification something can change at any given time.

Gödel's thought—comparable to Wittgenstein's later discussions of following a rule—is that if mathematics is constructed, one still has to decide, at each step of the construction, whether or not one has or has not constructed *correctly*. Poincaré's objection that the logicist “reduction” of mathematics requires mathematics itself to correct and titrate the carrying out of logic, would hold. In which case one has not really eliminated the idea of mathematical *truth*, as opposed to mere *logical correctness*.

The next remark of Gödel's [45] draws a conclusion for the idea of an *Urprinzip*:

*Remark (Philosophy):* One can say *complete agreement* [this written in English], but only in an improper sense *complete disagreement* [again written in English] [Note: that sounds comical]. This seems to indicate that every two things are somehow “one” (is this more than that every two things stand in some kind of connection?). <That> hangs together with the question whether the world is reducible to one or two symmetric *Urprinzipien*.

It is at this stage of *Max Phil IX* that we see Gödel's turn from Russell to Leibniz. It seems clear that he is sensitive to the threat of nonsense (chaos) and the possibility for a metaphysical grounding of his *Urprinzip*. Yet there is no solution posed in *Max Phil IX* with regard the difficulty of wholly false universes (or multiverses, or chaos). The limit point remains unclassified, and the problem of false objectives rears its head again, unresolved.

In the end, this suggests that Gödel's infinitary MRTJ foundered on the treatment of the base case—still in need of a clarification of “judgments of perception”—as well as the extrapolation to a unified view of truth. Gödel pursued investigation of these matters from a wide variety of angles in *Max Phil IX-X*, but without settling on a solution.

<sup>109</sup> Husserl offers an *Urprinzip* in (1913 I §24):

... every originary presentive intuition is a legitimizing source of cognition. ... everything originally offered us in “intuition” is to be accepted simply as what it is presented as being. . .

<sup>110</sup> On “chaos” see *Max Phil IX* [71].

In closing, we may remark that it is no accident that Gödel's philosophical notebooks bear such striking comparison with Wittgenstein's *Remarks on the Foundations of Mathematics* from 1940-1944.<sup>111</sup> The common points of origin are Russell's treatment of *n*th truth in *Principia* and his later (1940). Like Gödel, Wittgenstein had entered philosophy attempting to recast Russell's treatment of truth *via* acquaintance. Though he would recast this treatment in a very different direction from Gödel's, it should be clear that both Wittgenstein and Gödel picked up from Russell a demand for an analysis of truth in terms of belief and experience, and each sought a way to defend the unity of truth and explain the possibility of error.

For Gödel, this unity was to be conceptualized philosophically through dichotomous structuring of questions: truth must be conceived this way or that way, and then such and such will follow. For the later Wittgenstein, by contrast, the unity is to be earned bit by bit and shown, if at all, through investigation of the presuppositions and assumed pictures that make dichotomous structurings of questions about truth appear mandatory in a variety of particular situations. Gödel believed in universal reason as part of a necessary foundation for truth. Wittgenstein demanded that the notion of truth be recognizable immanently, within our human forms of life.

This great contrast in attitudes toward *the* fundamental problem of philosophy, revolves, however, about a shared philosophical inheritance. Just like the young Wittgenstein, Gödel found himself engaged in serious critical reflection on our uses of the verb "to perceive" in connection with truth and types when he pondered how to go on from *Principia*. Furthermore, Gödel did read Russell carefully, and had reached a view about the exchanges between Wittgenstein and Russell on this topic. He understood that Russell had travelled away from the theory of mathematical truth broached in *Principia* in his subsequent publications, at least partly under the influence of Wittgenstein.

What is fascinating is that Gödel recapitulates, in his *Max Phil* notebooks, many of the philosophical difficulties on which Russell and Wittgenstein—earlier and later—touched. After all, these problems naturally emerge from the MRTJ and Russell's later naturalistic and psychologistic writings. It is no accident that Gödel and Wittgenstein both turn to an investigation of our "feel" for everyday language, grammar, meaning, and truth, and write down so many "grammatical" remarks: they are responding to Russell's own sense of philosophy.

Gödel believed that Russell had, in the empiricist vein, always placed far too much emphasis upon sensation, focusing on examples such as "this is red" even in *Principia*, when presenting the MRTJ.<sup>112</sup> Russell had also wrongly promulgated a fact/ value dichotomy, and a kind of emotivism or conventionalism that Gödel regarded as ultimately authoritarian because merely conventionalist.<sup>113</sup> In all these respects, his

111 This is just as Wang, who probably derived the sense of it from Gödel himself, insisted (1987, 1996); cf. Floyd (2011).

112 Cf. *Principia*: p. 43 and Russell (1940: p. 26). Compare Wang's report (1996: p. 173, 5.4.4-5.4.5) that Gödel said that Russell made "drastic mistakes" in assuming that sense-perception is the only experience we can find by introspection.

113 Cf. Franks (2011). Toledo reports (2011: p. 200) Gödel saying:

The analytic philosophers try to make concepts clear by defining them in terms of primitive terms. But they don't attempt to make the primitive terms clear. Moreover, they

responses to Russell resembled Wittgenstein's. And yet: Gödel could not but applaud Russell's effort to resist the reduction of the notion of "truth" to that of "correctness" *via* a step into metaphysics.

## 6. Sources

### 6.1 Gödel's and Russell's philosophical writings

Which purely philosophical writings of Russell's did Gödel read? And when? In light of *Max Phil IX-X*, it is imperative to ask whether, when, and to what extent Russell's purely philosophical writings on belief, truth, consciousness, action and even emotion and politics penetrated Gödel's understanding—not only of Russell, but of his own philosophical thinking.

At the heart of our interpretation is the claim that Gödel was sufficiently aware of the MRTJ's structure to have tried to generalize it to an infinitary context in 1942-1943. We have emphasized as a key philosophical source chapter II of the (first and second edition's) Introduction to *Principia*. It is clear that despite the machinations in his theory of knowledge over the years after 1910, Russell retained the correspondence view, at least for empirical knowledge, ever after. To the extent that he did, this forms a remarkably bright line of affinity between Gödel and Russell over a number of decades.

We review here what we know of the current evidence concerning which among Russell's writings on truth Gödel would have had the opportunity to study. Those explicitly cited in (1944) we have already discussed (i.e., *Principia*, Russell's (1906a), (1906b), (1918), (1919a), and (1940)). It is not a bold conjecture, however, to suppose that Gödel would have occupied himself with Russell's purely philosophical arguments for the correspondence theory of truth. As Stadler points out (2001: p. 91), Gödel's general affinity with the objectivist and realist elements of a broader Austrian tradition must be taken seriously, and his reading of Russell would have fit into the specific Viennese period of his education.

As the editors of Schlick's collected papers have said, "to no other philosopher did Schlick devote so many lecture courses during his time in Vienna than to Russell" (Schlick (2008), p. 225). As we have said, Schlick discussed the MRTJ already in his (1910), and engaged with Russell's external world program in his (1925). After arriving in Vienna Schlick made sure, not only that *Principia* was available, but also Russell's other, more purely philosophical writings on truth. In 1923 Russell personally sent Schlick copies of *Our Knowledge of the External World* (1914b), *Introduction to Mathematical Philosophy* (1919a), and *Analysis of Mind* (1921). Schlick placed them in the Philosophical Institute library.<sup>114</sup> We have seen already that Gödel attended the 1925/1926 discussions of *Introduction to Mathematical Philosophy* in Vienna; but engagement with Russell continued. Russell's *Problems of Philosophy* (1912) was translated, probably at Schlick's behest, into German in 1926 and Schlick reviewed

---

take the wrong primitive terms, such as red etc., while the correct primitive terms would be object, relation, well, good, etc.

<sup>114</sup> Russell to Schlick, 20 February 1923, Schlick to Russell 18 March 1923, both at the McMaster Bertrand Russell Archives.

the translation (1927). Gödel attended Schlick's seminar in the winter semester of 1928/1929 on *The Problems of Philosophy*, as well as the subsequent semester's seminar Schlick gave on *The Analysis of Mind*. In the winter semester of 1929/30 Schlick lectured on Russell's *The Analysis of Matter* (1927a), giving a final seminar on *Introduction to Mathematical Philosophy* (1919a) in the summer semester of 1932. Gödel heard at least one of Schlick's lectures on "Knowledge" in 1934/1935, so that he would very probably have known Schlick's (1925), and so Schlick's discussion (in §10) of the MRTJ in (Russell (1912)).

As for Russell's other philosophical writings, the MRTJ was broached, without endorsement, in (1906/1907), a philosophical essay "On The Nature of Truth", where the "chief" reason for analyzing truth and belief as "multiple" relations is said to be paradoxes such as "the Liar" (p. 46). Given that it appeared in *The Proceedings of the Aristotelian Society*, Gödel may have known this essay in Princeton;<sup>115</sup> more likely still he could have studied its reprint and reworking in two chapters on truth in Russell's (1910b) *Philosophical Essays*—a book possibly available in Vienna and definitely in Princeton. (As we learn from Dawson (1997: ch. II) and records of borrowings of many books and journals written in English (and French), Gödel's linguistic knowledge was even in Vienna good enough for him to have read originals in English.) *Philosophical Essays* contains four essays on truth, including the one on "William James's Conception of Truth" (1908) where Russell stated that he considered the nature of truth to be "the fundamental question [of philosophy]".<sup>116</sup>

In addition, the articles Russell published in *The Monist* in 1906/7, 1914–1915 (the first few parts of (1913)) and 1918–19 (published later as (1918)) were available in full series at the University Library in Vienna, as well as in Princeton.

Between 3 October 1932 and 7 October 1932 Gödel apparently framed a plan to read Russell, writing a note in Gabelsberger shorthand that says "Russell: Philosophie des Geistes, Probleme der Philosophie". In English, this is "Philosophy of Mind" (presumably referring to the translation of Russell's *Analysis of Mind* by Grelling) and "Problems of Philosophy" (presumably referring to the translation of Russell's *The Problems of Philosophy* by Hertz). There is also another note in Gabelsberger on 7 October 1932 that says "Russell Philosophie der Materie (Philosophy of Matter = The Analysis of Matter) p. 177 § 80". As we have already seen (§5), he was then reading *The Analysis of Matter* (1927a) and on 7 October 1932 copied out an extract from the book concerned with Russell's calling the truths of mathematics, and not merely of logic, "tautologies". A version of the MRTJ was, as we have said, reworked in chapter XII, "Truth and Falsehood", of Russell's *Problems of Philosophy* (1912), available in the University Library in Vienna in English—and in German. Gödel borrowed the latter from the Vienna University Library on 28 November 1932, 2 January 1933, and again on 2 February 1933.

115 We have no evidence at present that *The Proceedings of the Aristotelian Society* or the Supplementary Volumes were available in Vienna before 1936.

116 The first two sections of Russell's (1906/7) were reprinted in (1910b) under the title "The Monistic Theory of Truth", sans Russell's remark about the Liar, which had occurred in section III of (1906/1907). Russell reworked that section III and published the new version as the final chapter of (1910b), "On the Nature of Truth and Falsehood".

On 3 November 1932 Gödel borrowed Grelling's German translation of *The Analysis of Mind* (1921) from the University Library in Vienna. Here Russell defended neutral monism and revisited the distinction between truth and falsehood, acknowledging explicitly (p. 272, n.1) that his views on the duality of truth and falsehood were owed to his "friend" Wittgenstein. (In response Wittgenstein would give a series of very critical lectures on the book at Cambridge, using it to develop some of his later ideas about consciousness and desire.<sup>117</sup>) A copy of Russell (1921) happens to have been purchased for the Institute for Advanced Study in 1940.

Given Gödel's entries in *Max Phil IX-X* on physics, emotion, will, action, perception, politics and history, it is worth remarking that many other works of Russell's—some of them of just the sort a beginning student of the subject like Gödel might have read—were available in the University and the Philosophical Libraries in Vienna (and later, in Princeton).

In 1930 Grelling translated *An Outline of Philosophy* (1927b) into German, and it was available before 1931 in the University Library in Vienna. Its chapter XIII, "Truth and Falsehood", introduces the notions of Truth and Falsehood, rehearsing, in the context of Russell's (then-recent) adoption of "neutral monism", what Russell calls "the formal problem of defining truth and falsehood, and deriving the objective reference of a proposition from the meanings of its component words" (p. 254ff). This presentation recapitulates Russell's treatment of negation and falsehood in terms of belief and the perception of complex universals. Not only does Russell argue explicitly against the idea that truth is a primarily a property of sentences, he insists on the complexity of belief and presents the MRTJ in terms of the idea that (p. 259) "a belief. . . interpreted narrowly, is a form of words related to an emotion of one of several kinds".

More generally—and consonant with the ambition and breadth both of Russell's and of Gödel's philosophical interests—while Gödel was a student the University Library in Vienna also contained a 1928 German translation by Grelling of *The ABC of Relativity* (1925), as well as numerous of Russell's political and historical writings. *Mysticism and Logic* (1917) was in the Vienna University Library (in a 1925 English edition) before 1931, as was the original essay that gave the book its title, reprinted as its chapter 1, "Mysticism and Logic" (1914a).

Of Russell's other writings *Sceptical Essays* (1928) and *Icarus, or the Future of Science* (1924) are of special interest for their sweep and characterization of the history of Western philosophy through the 1920s. Two themes on which Gödel dwells in *Max Phil IX* resonate, not only generally with the aims and era of the Vienna Circle, but with Russell's particular narratives: (1) the role of emotion in history and the prospects for managing it through exercise of the intellect and (2) an account of the broad sweep of the history of philosophy, told from the Russellian point of view: there was an old world, standing for synthesis and wholeness, and a new world, in which fragmentary, piecemeal, and scientifically specialized cooperative work would hold sway.

---

117 Cf. Monk (1990): p. 214, 291.

## 6.2 Gödel's correspondence with Russell

(1) In 1971 Kenneth Blackwell wrote to Gödel asking him to comment on his relationship with Russell, and in particular on Russell's remarks about Gödel in his *Autobiography* (1967, 1968, 1969: ch. 13). Russell had claimed that, while living in Princeton in 1943-1944, he "came to know Einstein fairly well" and "used to go to his house once a week to discuss with him and Gödel and Pauli", but that

. . . these discussions were in some ways disappointing, for, although all three of them were Jews and exiles and, in intention, cosmopolitans, I found that they all had a German bias towards metaphysics, and in spite of our utmost endeavors we never arrived at common premises from which to argue. Gödel turned out to be an unadulterated Platonist, and apparently believed that an eternal "not" was laid up in heaven, where virtuous logicians might hope to meet it thereafter.

While one cannot be sure whether the discussion took place before, during or after Gödel's submission of his manuscript of (1944) to the publisher,<sup>118</sup> on our reading it is no accident that it is the whole issue of "not"—that is, the distinction between truth and falsehood at ground level—to which Russell adverted in his recollection of meeting Gödel in Princeton. After all, for Russell, falsehood was *the* central conceptual difficulty to be faced in attacking, not merely the nature of "the simplest kind of *truth* and *falsehood*" at the primordial point, but in defeating the coherence theory of truth, the Leibnizian problem of whether a consistent fairy tale would be enough. And Gödel knew this. In fact, *Max Phil* IX begins with a marginal note: "From the possibility of negative being it follows that being is a relation of *potentiality* to the 'external world'".

Gödel's drafted reply to Kenneth Blackwell is interesting (2003a: p. 316-317). The anti-Continental flavor of Russell's remark was unfortunately compounded by his mistake in assuming Gödel was Jewish. To this Gödel replied that "for the sake of truth. . . I am not a Jew (even though I don't think ~~that~~ this question is of any importance)". He also noted that the passage "gives the wrong impression that I had many discussions with Russell, which was by no means the case. (I remember only one)". This means, of course, that there exists at least one conversation that they *did* have, and which Gödel remembered. Finally, and most importantly, Gödel wrote the following with regard to the purely philosophical charge (p. 317):

Concerning my "unadulterated" Platonism it ~~was/is~~ <is no more "unadulterated"> than Russell's own in 1921 when in <the> Introd. ~~was even~~ <he said>: [~~com the "not" with an animal when by negating~~] "...". At the time evidently Russell had met the "not" even in *this* world but later on <under the influence of Wittgenstein> he chose to overlook it.

Perhaps because of the reference to an "animal", as well as the phrase "Introd", the editors have changed "1921" in the above to "1919", Russell's *Introduction to Mathematical Philosophy*, to whose second edition (1919a) Gödel explicitly refers in (1944: p. 127) as he quotes his favorite remark from Russell that "logic is concerned with the real world just as truly as zoology, though with its more abstract and general features" (1919: p. 169). This editorial decision is understandable, for there was no 1921 edition of this book.

118 Dawson (1997: p. 163).

On this reading, Gödel's crossed out remark about "animals" is taken to refer, quite loosely, to Russell's famous remark about zoology.<sup>119</sup> (We discussed the *Max Phil* discussion of "logic as zoology" in §§3 and 5.2 above.) However, it is more likely that Gödel had in mind Russell's discussion of animal consciousness in the Introduction to (1921), where it is used to support his argument for neutral monism.

Be this as it may, it is clear from the letter draft that it was the influence of *Wittgenstein* on Russell that Gödel did not like, that he felt Wittgenstein had influenced Russell to go down the wrong path on the "not" "in this world". Animals, as well as Wittgenstein's views on particular judgments, are discussed at length in Russell's (1940). Russell even goes so far as to conjecture that animal hesitation accounts for the entering of disjunction into the framework (1940: pp. 84-85). Yet he once again separates from this the problem of negation, connecting it with the importance of distinguishing truth and falsehood at the ground, stating that this presents a deeper and more difficult set of issues.

(2) Gödel professed his philosophical motives to Russell directly, in a letter he wrote on the very day he sent in the final manuscript of (1944) to Schilpp (September 28, 1943). His revisions to (1944) had taken sufficiently long that Russell decided not to write a reply.<sup>120</sup> Gödel was quite disturbed by this, and wrote to Russell attempting to change his mind. But to no avail. Interestingly, although he duly mentions a technical error he had found in connection with *Principia* at \*89.16, the center of Gödel's plea to Russell concerns philosophy. He even expressed doubt, at this late stage, about whether he should publish the essay:

Dear Prof. Russell,

Prof. Schilpp has informed me that you don't intend to write a reply to my article in the Library of Living Philosophers. I regret that very much and am very sorry that this is apparently due to my sending in the manuscript so late. On the other hand I still hope that you may perhaps change your mind, since your decision seems to have been based on the wrong assumption that my article will not be controversial. This is by no means so. I am advocating in some respects the exact opposite of the development inaugurated by Wittgenstein and therefore suspect that many passages will contradict directly your present opinion. . . I doubt very much if my article should be published at all without a reply at least to the main items and you will perhaps agree with me after having read it.<sup>121</sup>

119 Wang (1987, p. 313) is puzzled by Gödel's footnote 3 in (1944 p. 127), in which, erroneously, Gödel states that the quote about zoology was not printed in later editions of (Russell 1919a). If our conjecture is correct, perhaps the mistake entered in with Gödel's mixing up Russell's (1919a) with (1921), which *did* have some changes made to it between printings (1921 and 1924), specifically at p. 111, in the discussion of "Introspection".

120 Instead, Russell concocted a general paragraph stating that he believed Gödel was essentially correct, and that he was himself no longer working in the foundations of mathematics. See Parsons (1990, p. 102), Schilpp and Russell (1944, p. 741).

121 For ease of readability, we have not reproduced the edits and redraftings of the letter, which are discussed in Goldfarb's introductory note (Gödel 2003: p. 207).

### 6.3 *Max Phil* notebooks IX-X

A brief overview of the *Max Phil* notebooks follows. They provide us with an unvarnished presentation of Gödel's manner of composition, and so provide a new opportunity for understanding certain aspects of Gödel's philosophy, though one that is, as has been noted, "difficult to survey".<sup>122</sup> So far we have worked synoptically, focussing on Gödel's engagement with Russellian themes so as to analyze the character of his initial move from mathematics toward pure philosophy. There are plenty of other materials of interest in *Max Phil* IX-X.<sup>123</sup> In Gödel (1944) and the *Max Phil* notebooks from 1943 onward one sees an increasing engagement with Leibniz and his philosophy of concepts, and other thinkers such as Plato, Aquinas, Leibniz and Carnap are discussed along the way. Yet in a real sense there remains persistent preoccupation with Russell.

*Max Phil* is a series of numbered notebooks in Gödel's *Nachlass*, I-XV (with XIII apparently lost), written by Gödel in German Gabelsberger shorthand. Dated in the years 1938 to 1955, these notebooks, amounting to some 1500 pages, comprise the written record of Gödel's philosophical explorations during the period. What will soon be available, thanks to the French Project directed by Gabriella Crocco, are notebooks IX-X, spanning the period November 1942 through January 1944. Of special interest with respect to Gödel's reaction to Russell are his remarks on truth, definition, psychology, language, perception, understanding and grammar in these two *Max Phil* Notebooks, from which we have drawn. Here the progress of Russell's philosophy from 1910 through 1940 is Gödel's central concern, his serious turn to Leibniz coming later.<sup>124</sup>

Gödel's (1944) was solicited by Paul Schilpp in a letter of 18 November 1942 to contribute to the volume on Russell in the Library of Living Philosophers series (Schilpp 1944).<sup>125</sup> That very day Gödel began the *Max Phil* IX notebook. Whereas his previous six notebooks III-VIII were paginated continuously p. 1-680, he started IX anew at page 1, indicating a fresh beginning. Gödel completed a draft manuscript of the Russell essay in just six months, submitted it on 17 May 1943, and then a final manuscript on 28 September 1943. *Max Phil* IX is dated to 11 March 1943, and *Max Phil* X has the dates 12 March 1943 to 27 January 1944. Hence Gödel was working on the Russell essay while writing *Max Phil* IX and much of *Max Phil* X.

Because of their working nature, the *Max Phil* notebooks must be used with care. A comparison with Wittgenstein's nearly contemporaneous writings on Russell is perhaps useful in this respect. Working notebooks never intended for publication provide neither a report nor a settled touchstone for assessing a philosopher's views. Yet they have their place, if one seeks a glance over the thinker's shoulder to get a feel

122 Dawson and Dawson (2010, p. 23).

123 Much could be said, for example, about Gödel's remarks on constructivism and intuitionism. For months up to this time Gödel had been working on the independence of the Axiom of Choice in intuitionistic type theory, this being a not unnatural follow-up to his lecture (1933a) and his work on *L*. However, he had become quite dissatisfied, and we see him focus with renewed energy and enthusiasm on philosophical issues; cf. Dawson (1997, p. 162f).

124 See van Atten and Kennedy (2003), Crocco (2006, 2012).

125 Gödel (2003b, p. 217).

for the intellectual sensibility at work, as well as a sense of the philosopher's aims and standards of argumentation.

Gödel's notebooks, especially wide-ranging, ambitious, and yet tentative during the period we are discussing, are, as Crocco has said, a "kaleidoscopic" parade of thought-seeds, argument sketches, and observations.<sup>126</sup> Co-editor Engelen has pointed out that the very name "*Max Phil*" derives from "Maxims", in the tradition of Goethe, whose "Maxims and Reflections" were collected after his death.<sup>127</sup> While Russell disparaged Wittgenstein's later writing style as "aphoristic" (1959: p. 93), the *Max Phil* notebooks, as a collection of remarks, sentences, analogies and examples has a similar form, consisting as it does of brief "remarks". If anything, Gödel attends far less to the nature of philosophical argumentation than does Wittgenstein. With Gödel one has an impression of rougher, more exploratory work, less orchestrated and sustained, less ambivalent, focused on extracting ideas from Russell and more deeply steeped in mathematics, though ranging widely. Examples from the history of philosophy, mathematics, music, logic and politics, grammatical, philosophical, and theological remarks are interwoven, showing us in practice Gödel's sense of what philosophy is and ought to be. As Gödel wrote (*Max Phil* IX, (78-79)),

**Remark** (*Philosophy*): My work in connection with *philosophy* should consist in an analysis of the highest concepts (logical and *psychological*); that is, what in the end is to be done is to write down a list of these concepts and the possible *axioms*, *theorems*, and *definitions* to be considered for them (obviously along with application to the empirically given reality). In order to be able to do this, one must first acquire, through (half-understood) philosophical readings [note: and the writing down of philosophical "*Remarks*"] a feel for what one can accept. On the other hand the understanding of an *axiomatics* will increase the understanding of the philosophical writers [thus the interaction of "top" and "bottom", where the right relation is important].

A substitute for the reading of philosophers is reading some good books with precise analyses [two words are stricken out here, no longer readable], learning of language [*Hebrew, Chinese, Greek*] and the precise *definition* of words and concepts that occur.

Like Wittgenstein, Gödel returns often to issues of language and meaning. After all, to make out his analogy with sense-perception for concepts, he sees that he will have to have an account of what it is to grasp meaning through reflection, when not every question is decidable in a formalism. Like Wittgenstein, he refuses to accept either conventionalism about lexical meaning, or Russell's later psychologistic reductionism about our grasp of meaning. Unlike Wittgenstein, in *Max Phil* IX-X Gödel tends to sort his remarks into different thematic areas by label, and he does not obsessively focus on argumentation and what it reveals about the nature of philosophy. He simply sets down ideas, and attempts to lay out dichotomous choices.

In *Max Phil* IX-X there are 100 "Grammatical", 117 "Philosophical", and 31 other remarks, including unlabelled ones: (9) on history, politics, and metaphysics, as well as remarks labelled "Theological" (10), "Mathematical" (1), "Physics" (4), and "Psychological" (6), the latter containing some direct reflections by Gödel on his own development and life. However, remarks labelled "Philosophical" often contain remarks about physics or grammar in them, so divisions are arguable.

<sup>126</sup> Cf. Crocco (2015).

<sup>127</sup> Cf. the Editors' note in the current draft of *Max Phil* IX.

In general we see an increasing number of remarks labelled “Grammatical” as the notebooks progress, although there are internal patterns within these. A series of briefer “Philosophical” remarks will culminate in a longer entry, sometimes containing a numbered encapsulation of points or theses as Gödel seeks to formulate definitions, after which focus is shifted back to a series of “Grammatical” remarks, which then lead back to a train of “Philosophical” remarks. Often what is labelled a “Grammatical” remark looks like what we would straightforwardly call a “Philosophical” or a “Logical” one (e.g., when Gödel discusses Russell’s theory of descriptions, or his treatment of logical fictions (*Max Phil* IX [49]), though not always. And often there is interpenetration, as when an analogy between physics and sociology is treated under the auspices of a “Philosophical” remark (*Max Phil* X [4]), or when Gödel simply writes down “Philosophy=Religion” (*Max Phil* X [4]). It is striking how much faith Gödel places in the return to immediate reflection on aspects of language that he takes the later Russell to have neglected: the “feel” of a language (*Max Phil* IX [3]); the nasal and consonantal sound of phonemes (*Max Phil* IX [57]); the emotional, historical, and political effects of particular songs (the Marseillaise, Lincoln and Liberty (*Max Phil* IX [4]). Here again is an intersection with Wittgenstein’s way of working, however orthogonal it is to Gödel’s in aim: the philosopher dips back into grammar, history, and ordinary language, not because they are sacrosanct, but because they hold clues to fresh ways of framing questions that mathematics and philosophy can distort and cover up.

## Acknowledgment

Thanks are due to John W. Dawson Jr., Juliette Kennedy and Marcia Tucker for their work in making Gödel’s Nachlass available to scholars.

This is a much expanded version of a lecture delivered by the second author at the Gödel ANR conference at Aix-en-Provence in July 2013. Our thanks are due to the organizers of the conference for their work in editing and transcribing the *Max Phil* notebooks, a project first broached by Cheryl Dawson as well to the several participants who offered suggestions, e.g. John Dawson, Göran Sundholm, and Mark van Atten. We are also grateful to Kenneth Blackwell, Gabriella Crocco, Arlene Duncan, and Rick Stapleton (Bertrand Russell Research Centre, McMaster University), Alexandra Metz (Vienna Philosophical Institute Library), Eva-Maria Engelen, Juliette Kennedy, Montgomery Link, Jean-Philippe Narboux, Sanford Shieh and Jouko Väänänen for helpful suggestions as to historical and philosophical matters during the essay’s writing. We are, finally, indebted to Hilary Putnam and Wilfried Sieg for comments on a late draft of the paper, and to Jared Henderson for his lightning production of the Word file.

## References

- AWODEY, Steve and CARUS, André W., (2010) "Gödel and Carnap", in Feferman, Parsons, Simpson, eds., (2010) pp. 252-273.
- BEANEY, Michael ed. (2013) *The Oxford Handbook of The History of Analytic Philosophy*, Oxford: Oxford University Press.
- BERNAYS, Michael ed. (2013) *The Oxford Handbook of The History of Analytic Philosophy*, Oxford: Oxford University Press.
- BERNAYS, Paul (1946) "Review of 'Russell's Mathematical Logic' by Kurt Gödel", *The Journal of Symbolic Logic*, 11(3), pp. 75-79.
- BURGESS, John P., (2014) "Intuitions of Three Kinds in Gödel's Views on the Continuum", in Kennedy, Juliette, ed. (2014), pp. 11-31.
- CAREY, Rosalind (2003) "The Development of Russell's Diagrams for Judgment", *Russell: The Journal of Bertrand Russell Studies*, 23, pp. 27-41.
- CAREY, Rosalind (2007) *Russell and Wittgenstein on the Nature of Judgment*, New York: Continuum International Publishing Group.
- CARNAP, Rudolf (1934) *Die Logische Syntax der Sprache*, Vienna: Springer. Translated into English by Amethe Smeaton as *The Logical Syntax of Language*, Kegan Paul, Trench, Trubner, 1937.
- CARNAP, Rudolf (2000) *Untersuchungen zur Allgemeinen Axiomatik*. Edited from unpublished manuscript by T. Bonk and J. Mosterin. Darmstadt: Wissenschaftliche Buchgesellschaft.
- CROCCO, Gabriella (2006) "Gödel on Concepts", *History and Philosophy of Logic*, 27, pp. 171-191.
- CROCCO, Gabriella (2012) "Gödel, Leibniz, and 'Russell's Mathematical Logic'", in Krömer, Ralf and Chin-Drian, Yannick, eds., *New Essays on Leibniz Reception: In Science and Philosophy of Science 1800-2000*, Basel: Springer, pp. 217-256.
- CROCCO, Gabriella (2015) "Sinn/Bedeutung and Intension/Extension in Gödel's *Max PhilIX*", in Crocco, Gabriella and Engelen, Eva-Maria, eds., *Kurt Gödel: Philosopher-Scientist*, Aix-en-Provence: Presses Universitaires de Provence.
- DAVIS, Martin (2005) "What Did Gödel Believe and When Did He Believe It?" *The Bulletin of Symbolic Logic*, 11, pp. 194-206; reprinted in Feferman *et al.*, eds., (2010), pp. 229-241.
- DAWSON, Cheryl A. and DAWSON, John W. Jr., (2010) "Future Tasks for Gödel Scholars", in Feferman, Solomon, Parsons, Charles and Simpson, Stephen G., eds., *Kurt Gödel: Essays for His Centennial*, Lecture Notes in Logic, vol. 33, Cambridge: U.K./New York: Association for Symbolic Logic/Cambridge University Press, pp. 21-44.
- DAWSON, John W., Jr., (1997) *Logical Dilemmas: The Life and Work of Kurt Gödel*, Wellesley: A. K. Peters.
- DEDEKIND, Richard (1888) *Was sind und was sollen die Zahlen?*, Braunschweig: Vieweg. Reprinted in Ewald William, ed., *From Kant to Hilbert: A Sourcebook in the Foundations of Mathematics*, vol. 2, Oxford: Oxford University Press, pp. 790-833.
- DREBEN, Burton and FLOYD, Juliet (1991) "Tautology: How Not to Use a Word", *Synthese*, 87(1), pp. 23-50.
- FEFERMAN, Solomon (1962) "Transfinite Recursive Progressions of Axiomatic Theories", *The Journal of Symbolic Logic*, 27, pp. 259-316.
- FEFERMAN, Solomon (1964) "Systems of Predicative Analysis", *The Journal of Symbolic Logic*, 29, pp. 1-30.

- FEFERMAN, Solomon (1987) “Infinity in Mathematics: Is Cantor necessary? (Conclusion)”, *L’infinito nella Scienza, Roma, Istituto della Enciclopedia Italiana*, pp. 151-209. Reprinted in Feferman (1998), pp. 229-248.
- FEFERMAN, Solomon (1998) *In the Light of Logic*, New York: Oxford University Press.
- FEFERMAN, Solomon, PARSONS, Charles, and SIMPSON Steven G., eds., (2010) *Kurt Gödel: Essays for His Centennial*, vol. 33, Lecture Notes in Logic, Cambridge: Association for Symbolic Logic/Cambridge University Press.
- FLOYD, Juliet (1995) “Wittgenstein, Gödel, and the Trisection of the Angle”, in Hintikka, Jaakko, ed., *From Dedekind to Gödel: Essays on the Development of the Foundations of Mathematics*, Synthese Library, Dordrecht: Kluwer Academic, pp. 373-426.
- FLOYD, Juliet (2001) “Prose versus Proof: Wittgenstein on Gödel, Tarski, and Truth”, *Philosophia Mathematica*, 3(9), pp. 901-928.
- FLOYD, Juliet (2011) “Wang and Wittgenstein”, in Parsons, Charles and Link, Montgomery, eds., *Hao Wang, Logician and Philosopher*, Texts in Philosophy, London: College Publications, pp. 145-194.
- FLOYD, Juliet (2012a) “Wittgenstein, Carnap, and Turing: Contrasting Notions of Analysis”, in Wagner, Pierre, ed., *Carnap’s Ideal of Explication and Naturalism*, History of Analytic Philosophy, Basingstoke: Palgrave Macmillan, pp. 34-46.
- FLOYD, Juliet (2012b) “Wittgenstein’s Diagonal Argument: A Variation of Cantor and Turing”, in Dybjer, Peter, Lindström, Per, Palmgren, Erik and Sundholm, Göran, eds., *Epistemology versus Ontology: Essays on the Philosophy and Foundations of Mathematics in Honour of Per Martin-Löf*, vol. 27, Dordrecht: Springer Netherlands, pp. 25-44.
- FLOYD, Juliet (2013a) “Turing, Wittgenstein and Types: Philosophical Aspects of Turing’s ‘The Reform of Mathematical Notation and Phraseology’ (1944-1945)”, in Cooper, Barry S., and van Leeuwen, Jan, eds., *Alan Turing—His Work and Impact*, Amsterdam: Elsevier Science, pp. 250-253.
- FLOYD, Juliet (2013b) “The Varieties of Rigorous Experience”, in Beaney, ed. (2013) pp. 1003-1042.
- FLOYD, Juliet and PUTNAM, Hilary (2012) “Wittgenstein’s ‘Notorious’ Paragraph about the Gödel Theorem: Recent Discussions”, in Putnam (2012), pp. 458-481.
- FRANKS, Curtis (2011) “Stanley Tennenbaum’s Socrates”, in Kennedy and Kossak, eds., (2011), pp. 208-225.
- FREGE, Gottlob (1893). *Grundgesetze der Arithmetik, Begriffsschriftlich abgeleitet*, vol. 1, Jena: Hermann Pohle. Reprinted Hildesheim, Olms (1962). English translation (2013), *Basic Laws of Arithmetic Derived Using Concept-Script*, vols I,II, trans. and ed. by P. A. Ebert and Marcus Rossberg, Oxford: Oxford University Press.
- FREGE, Gottlob (1918) “Der Gedanke”, in *Beiträge zur Philosophie des deutschen Idealismus*, pp. 58-77. Translated into English as “Thought” by Peter Geach and R.H. Stoothoff in *Gottlob Frege, Collected Papers on Mathematics, Logic, and Philosophy*, McGuinness, Brian, ed., New York and Oxford: Oxford University Press, 1984, pp. 351-372; also in *The Frege Reader*, Beaney, Michael, ed., Oxford: Blackwell, 1997, pp. 325-345.
- FROST-ARNOLD, Greg (2013) *Carnap, Tarski and Quine at Harvard: Conversations on Logic, Mathematics, and Science*, Chicago, IL: Open Court.
- GÖDEL, Kurt (1930) “Die Vollständigkeit der Axiome des logischen Funktionenkalküls”, *Monatshefte für Mathematik und Physik*, pp. 349-360. Reprinted and translated in Gödel (1986), pp. 102-103.

- GÖDEL, Kurt (1931) “Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I”, *Monatshefte für Mathematik und Physik*, 38, pp. 173-198. Reprinted and translated in Gödel (1986), pp. 144-195.
- GÖDEL, Kurt (1932) “Über Vollständigkeit und Widerspruchsfreiheit”, *Ergebnisse eines mathematischen Kolloquiums*, vol. 3, pp. 12-3. Reprinted and translated in Gödel (1986), p. 234.
- GÖDEL, Kurt (1933o) “The Present Situation in the Foundations of Mathematics”, Handwritten text for an invited lecture, in Gödel (1995), pp. 45-53.
- GÖDEL, Kurt (1938) “The Consistency of the Axiom of Choice and of the Generalized Continuum-Hypothesis”, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 24, pp. 556-557. Reprinted in Gödel (1990), pp. 234-237.
- GÖDEL, Kurt (1939a) “Consistency-proof for the Generalized Continuum-Hypothesis”, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 25, pp. 220-224. Reprinted in Gödel (1990), pp. 28-32.
- GÖDEL, Kurt (1939b) “Vortrag Göttingen”, text and translation in Gödel (1995), pp. 126-155.
- GÖDEL, Kurt (1940) *The Consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis with the Axioms of Set Theory*, vol. 3 of *Annals of Mathematics Studies*, Princeton: Princeton University Press. Reprinted in Gödel (1990), pp. 33-101.
- GÖDEL, Kurt (1940a) “Lecture [on the] consistency [of the] continuum hypothesis (Brown University)”, in Gödel (1995), pp. 175-185.
- GÖDEL, Kurt (1942-1943/unpublished), *Max Phil IX: 18 November 1942 - 11 March 1943*, notebook at the Institute for Advanced Study, Princeton University. Transcription from Gabelsberger shorthand to German with editing under the auspices of the project ANR-09-BLA-0313, edited by Gabriella Crocco, *et al.*, transcription from Gabelsberger shorthand to German by Englen Eva-Maria based on an earlier draft by Cheryl Dawson. English trans. by Juliet Floyd; the draft transcription relied upon for the present paper is available from the trans. at [jfloyd@bu.edu](mailto:jfloyd@bu.edu).
- GÖDEL, Kurt (1943-1944/unpublished), *Max Phil X: 12 March 1943 - 27 January 1944*, notebook at the Institute for Advanced Study, Princeton University. Transcription from Gabelsberger shorthand to German with editing under the auspices of the project ANR-09-BLA-0313, edited by Gabriella Crocco, *et al.*, transcription from Gabelsberger to German by Englen Eva-Maria based on an earlier work by Cheryl Dawson and Robin Rollinger, introduction and annotations by G. Crocco. English trans. by Juliet Floyd; the draft transcription relied upon for the present paper is available from the trans. at [jfloyd@bu.edu](mailto:jfloyd@bu.edu).
- GÖDEL Kurt (1944) “Russell’s Mathematical Logic”, in Schilpp, Paul A., ed., *The Philosophy of Bertrand Russell*, vol. 5 of Library of Living Philosophers, Evanston: Northwestern University Press, pp. 123-153. Reprinted in Gödel (1990), pp. 119-141.
- GÖDEL, Kurt (1946) “Remarks Before the Princeton Bicentennial Conference on Problems in Mathematics”. Reprinted in Gödel (1990), pp. 150-153.
- GÖDEL, Kurt (1947) “What is Cantor’s continuum problem?” *American Mathematical Monthly*, pp. 515-525. Reprinted in Gödel (1990), pp. 176-187.
- GÖDEL, Kurt (1958) “Über eine bisher noch nicht benutzte Erweiterung des finiten Standpunktes”, *Dialectica*, pp. 280-287. Reprinted and translated in Gödel (1990), pp. 240-251.
- GÖDEL, Kurt (1970a) “Some Considerations Leading to the Probable Conclusion that the True Power of the Continuum is  $\aleph_2$ ”, in Gödel (1995), pp. 420-422.
- GÖDEL, Kurt (1970b) “A Proof of Cantor’s Continuum Hypothesis from a Highly Plausible Axiom About Orders of Growth”, in (Gödel 1995), pp. 422-423.
- GÖDEL, Kurt (1986) *Collected Works, Volume I: Publications 1929-1936*, Oxford: Clarendon Press, ed. by Solomon Feferman *et al.*

- GÖDEL, Kurt (1990) *Collected Works, Volume II: Publications 1938-1974*, New York: Oxford University Press, ed. by Solomon Feferman *et al.*
- GÖDEL Kurt (1995) *Collected Works, Volume III: Unpublished Essays and Lectures*, New York: Oxford University Press, ed. by Solomon Feferman *et al.*
- GÖDEL, Kurt (2003a) *Collected Works, Volume IV: Correspondence A-G*, Oxford: Clarendon Press, ed., by Solomon Feferman, John W. Dawson Jr., *et al.*
- GÖDEL, Kurt (2003b) *Collected Works, Volume V: Correspondence H-Z*, Oxford: Clarendon Press, ed., by Solomon Feferman, John W. Dawson Jr., *et al.*
- GÖDEL, Kurt [unpublished], *Max Phil IX: 18 November 1942 - 11 March 1943*, transcription by Engelen Eva-Maria based on earlier draft by Dawson Cheryl.
- GÖDEL, Kurt (2015) *Max Phil X*, ed. by Crocco, Gabriella *et al.*, transcription by Engelen, Eva-Maria, based on earlier work by Dawson, Cheryl and Rollinger, Robin, introduction and annotations by G. Crocco, Aix-en-Provence: PUP.
- GOLDFARB, Warren (1989) "Russell's reasons for Ramification", *Rereading Russell, Essays in Bertrand Russell's Metaphysics and Epistemology*, Savage C. W. and Anderson C. A., eds., Minneapolis, MN: University of Minnesota Press, pp. 24-40.
- GOLDFARB, Warren (2005) "On Gödel's way in: The influence of Rudolf Carnap", *The Bulletin of Symbolic Logic*, 11, pp. 185-193; reprinted in Feferman *et al.*, (2010), pp. 242-251.
- GRIFFIN, Nicholas (1985) "Wittgenstein's Criticism of Russell's Theory of Judgment", *Russell*, 5, pp. 132-145.
- HARDY, G.H., (1929) "Mathematical Proof", *Mind*, 38(149), pp. 1-25.
- HAZEN, Allen P., (2000) "A "Constructive" Proper Extension of Ramified Typed Theory (The Logic of *Principia Mathematica*, Second Edition, Appendix B)", in Link, Godehard, ed., *One Hundred Years of Russell's Paradox: Mathematics, Logic, Philosophy*, deGruyter Series in Logic and Its Applications 6, Berlin and New York: de Gruyter, pp. 449-480.
- HAZEN, Allen P. and DAVOREN Jen M., (2000) "Russell's 1925 Logic", *Australasian Journal of Philosophy*, 78(4), pp. 534-556.
- HILBERT, David (1926) "Über das Unendliche", *Mathematische Annalen*, 95, pp. 161-190. Translated into French by André Weil in *Acta Mathematica* 48:91-122, 1926; translated into English in van Heijenoort, ed., (1967), pp. 367-392.
- HILBERT, David and W. ACKERMANN (1928) *Grundzüge der theoretischen Logik*, Berlin: Springer.
- HUSSERL, Edmund (1913) *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie, Erstes Buch: Allgemeine Einführung in die reine Phänomenologie*, Halle: Max Niemeyer Verlag. Translated into English by F. Kersten as *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy—First Book: General Introduction to a Pure Phenomenology*, Nijhoff: The Hague, 1982.
- HYLTON, Peter (1990) *Russell, Idealism, and the Emergence of Analytic Philosophy*, New York: Oxford University Press.
- JAMES, William (1904) "Does 'Consciousness' Exist?" *Journal of Philosophy, Psychology, and Scientific Methods*, (18), pp. 477-491.
- KANAMORI, Akihiro (2007) "Gödel and Set Theory", *The Bulletin of Symbolic Logic*, 13, pp. 153-188, reprinted in Feferman *et al.* (2010), pp. 145-180.
- KANAMORI, Akihiro (2012) "In Praise of Replacement", *The Bulletin of Symbolic Logic*, 18, pp. 45-90.
- KENNEDY, Juliette (2013) "On Formalism Freeness: Implementing Gödel's 1946 Princeton Bicentennial Lecture", *The Bulletin of Symbolic Logic*, 19, pp. 351-393.

- KENNEDY, Juliette (2014) “Gödel’s 1946 Princeton Bicentennial Lecture: An Appreciation”, in Kennedy, Juliette, ed., (2014) pp. 109-130.
- KENNEDY, Juliette, ed., (2014) *Interpreting Gödel*, Cambridge: Cambridge University Press.
- KENNEDY, Juliette (2015) “On the ‘Logic Without Borders’ Point of View”, in Hirvonen Asa, Kontinen, Juha Kossak, Roman and Villaveces, Andres, eds., *Logic without Borders: Essays on Set Theory, Model Theory, Philosophical Logic, and Philosophy of Mathematics*, Berlin: DeGruyter.
- KENNEDY, Juliette and KOSSAK, Roman, eds., (2011) *Set Theory, Arithmetic and Foundations of Mathematics: Theorems, Philosophies*, vol. 36 of Lecture Notes in Logic, New York: Cambridge University Press, pp. 208-225.
- KENNEDY, Juliette and VAN ATTEN, Mark (2009) “Gödel’s modernism: On set-theoretic incompleteness, revisited”, in Lindström Palmgren, S., E., Segerberg, K. and Stoltenberg-Hansen, V., eds., *Logicism, Intuitionism, and Formalism: What has become of them?*, New York: Springer, pp. 303-356.
- KREISEL, Georg (1980) “Kurt Gödel, 28 April 1906-1914 January 1978”, *Biographical Memoirs of the Fellows of the Royal Society*, 26, pp. 149-224. Corrections, 27 (1981), p. 697 and 28 (1982), p. 718.
- KREISEL, Georg (1983) “Einige Erläuterungen zu Wittgensteins Kummer mit Hilbert und Gödel”, in Weingartner, Paul A. and Czermak, Johannes, eds., *Epistemology and Philosophy of Science. Proceedings of the 7th International Wittgenstein Symposium*, Vienna: Holder-Pichler-Tempsky Verlag, pp. 295-303.
- LANDINI, Gregory (1998) *Russell’s Hidden Substitutional Theory*, Oxford: Oxford University Press.
- LANDINI, Gregory (2007) *Wittgenstein’s Apprenticeship with Russell*, Cambridge: Cambridge University Press.
- LEVINE, James (2009) “From Moore to Peano to Watson: The Mathematical Roots of Russell’s Naturalism and Behaviorism”, *The Baltic International Yearbook of Cognition, Logic, and Communication*, vol. 4, pp. 1-126.
- LEVINE, James (2013) “Principia Mathematica, the Multiple-Relation Theory of Judgment and Molecular Facts”, in Griffin, Nicholas and Linsky, Bernard, eds., *The Palgrave Centenary Companion to Principia Mathematica*, Basingstoke: Palgrave Macmillan, pp. 247-304.
- LINSKY, Bernard (1999) *Russell’s Metaphysical Logic*, CSLI Lecture Notes, Number 101, Stanford: CA, CSLI Publishing.
- MCGUINNESS, Brian (2008) *Wittgenstein in Cambridge: Letters and Documents 1911-1951*, Malden, MA, Oxford: Blackwell, 4th edition. All letters reprinted in Wittgenstein (2005).
- MONK, Ray (1990) *Ludwig Wittgenstein: The Duty of Genius*, New York/London: Free Press/Jonathan Cape.
- MONK, Ray (2000) *Bertrand Russell, 1921-1970: the Ghost of Madness*, New York/London: Jonathan Cape/Free Press.
- MOORE, G.E., (1899) “The Nature of Judgment”, *Mind*, New Series, 8(30), pp. 176-193.
- NEALE, Stephen (1995) “The Philosophical Significance of Gödel’s Slingshot”, *Mind*, 104(416), pp. 761-825.
- OBERDAN, Thomas (2013) “Moritz Schlick”, in *The Stanford Encyclopedia of Philosophy*. E. N. Zalta ed. May 2013, at <http://plato.stanford.edu/entries/schlick/>.
- PARSONS, Charles (1990) “Introductory note to Gödel 1944”, in Gödel (1990) pp. 102-118. Reprinted with a postscript in Parsons (2014), pp. 103-126.

- PARSONS, Charles (1995) "Platonism and mathematical intuition in Kurt Gödel's thought", *The Bulletin of Symbolic Logic*, 1, pp. 44-74. Reprinted in Ferferman *et al.* (2010), pp. 326-355 and reprinted with a postscript in Parsons (2014), pp. 153-198.
- PARSONS, Charles (2010) "Gödel and Philosophical Idealism", *Philosophia Mathematica*, 18(2), pp. 166-192.
- PARSONS, Charles (2014) *Philosophy of Mathematics in the Twentieth Century* Cambridge, MA: Harvard University Press.
- PEARS, David (1977) "The Relation Between Wittgenstein's Picture Theory of Propositions and Russell's Theories of Judgment", *The Philosophical Review*, 86(2), pp. 177-196.
- PINCOCK, Christopher (2008) "Russell's Last (and Best) Multiple-Relation Theory of Judgment", *Mind*, 117(465), pp. 107-139.
- POTTER, Michael (2009) *Wittgenstein's Notes on Logic*, New York: Oxford University Press.
- PUTNAM, Hilary (1985) "Reflexive Reflections", *Erkenntnis*, 22(1/3), pp. 143-153. Reprinted in Hilary Putnam, (1994) *Words and Life*, Harvard: University Press, pp. 416-427.
- PUTNAM, Hilary (1999) *The Threefold Cord: Mind, Body, and World*, New York: Columbia University Press.
- PUTNAM, Hilary (2012) *Philosophy in an Age of Science: Physics, Mathematics, and Skepticism*, Cambridge, Mass.: Harvard University Press. Edited by Mario de Caro and David MacArthur.
- PUTNAM, Hilary (2013) "Comments on Travis and McDowell", in Baghramian, Maria, ed., *Reading Putnam*, New York: Routledge, pp. 347-358.
- QUINE, Willard van Orman (1985) *The Time of My Life: An Autobiography*, Cambridge: MIT Press.
- QUINE, Willard van Orman (1992) *Pursuit of Truth*, Cambridge: MA, Harvard University Press.
- RAMSEY, Frank Plumpton (1923) "Review of *Tractatus Logico-Philosophicus* by Ludwig Wittgenstein", *Mind*, 32(128), pp. 465-478.
- RAMSEY, Frank Plumpton (1927) "Facts and Propositions", *Aristotelian Society Supplementary Volume*, vol. 7, pp. 153-70. Reprinted in *Proceedings of the Aristotelian Society*, The Virtual Issue n°. 1, 2013.
- RICKETTS, Thomas (1996) "Pictures, Logic, and the Limits of Sense in Wittgenstein's *Tractatus*", in Sluga, Hans and Stern, David eds., *The Cambridge Companion to Wittgenstein*, New York: Cambridge University Press, pp. 59-99.
- RICKETTS, Thomas (2001) "Truth and Propositional Unity in Early Russell", in Floyd, Juliet and Shieh, Sanford, eds., *Future Pasts: Reflections on the History and Nature of Analytic Philosophy*, New York: Oxford University Press, pp. 101-121.
- ROYCE, Josiah (1900) *The World and the Individual*, New York: The Macmillan Co.
- RUSSELL, Bertrand (1900) *A Critical Exposition of the Philosophy of Leibniz, with an Appendix of Leading Passages*, London: George Allen and Unwin. New edition published 1937.
- RUSSELL, Bertrand (1903) *The Principles of Mathematics*, London: Allen and Unwin.
- RUSSELL, Bertrand (1904) "Meinong's Theory of Complexes and Assumptions", *Mind*, 13, pp. 204-219, 336-354, 509-524. Reprinted in Russell (1994), ch. 17.
- RUSSELL, Bertrand (1905) "Necessity and Possibility", Read to the Oxford Philosophical Society on 22 October 1905. Published in Russell (1994), ch. 22.
- RUSSELL, Bertrand (1906a) "Les paradoxes de la logique", *Revue de Métaphysique et de Morale*, 14, pp. 627-650. Translation as "On 'Insolubilia' and Their Solution by Symbolic Logic" in Russell (1973) pp. 190-214. Reprinted in Russell (2014), ch. 9.

- RUSSELL, Bertrand (1906b) "On Some Difficulties in the Theory of Transfinite Numbers and Order Types", *Proceedings of the London Mathematical Society*, vol. 2(4), pp. 29-53. Reprinted in Russell (1973), pp. 135-164, and Russell (2014), ch. 2, as Russell (1905-1906).
- RUSSELL, Bertrand (1906/1907) "On the Nature of Truth", *Proceedings of the Aristotelian Society*, vol. 7, pp. 28-49. Reprinted in Russell (1994), ch. 14 as "[1907]".
- RUSSELL, Bertrand (1908) "Transatlantic Truth", *The Albany Review*, 2, pp. 349-353. Reprinted under the title "William James's Conception of Truth" in Russell (1910b), ch. 5 and in Russell (2014), ch. 16.
- RUSSELL, Bertrand (1910a) "On the Nature of Truth and Falsehood", *Philosophical Essays*, London, Longmans, Green, pp. 170-185. Reprinted in Russell (1992), ch. 12.
- RUSSELL, Bertrand (1910b) *Philosophical Essays*, London, Longman, Green. Republished in Routledge Classics, London and New York, 2009.
- RUSSELL, Bertrand (1912) *The Problems of Philosophy*, London, Williams and Norgate. Translated into German by Paul Hertz as *Die Probleme der Philosophie*, Weltkreis-Verlag: Vienna, 1926. Reprinted, Hackett Publishing Company, Indianapolis 1990.
- RUSSELL, Bertrand (1913) *Theory of Knowledge: The 1913 Manuscript*, vol. 7 of *The Collected Papers of Bertrand Russell*, London, Allen and Unwin, ed. by Elizabeth Ramsden Eames, (1984). An unpublished manuscript first discovered in 1967. Parts I, II (pp. 1-142) were published by Russell in a series of six articles in *The Monist*, 1914 (vol. 24, 1, pp. 1-16, 2, pp. 161-87, 3, pp. 435-453, 4, pp. 582-593) and 1915 (vol. 25, 1, pp. 28-44, 2, pp. 212-233). The remaining pp. 143-350 of the manuscript was first published by Routledge in 1984 in this volume.
- RUSSELL, Bertrand (1914a) "Mysticism and Logic", *The Hibbert Journal*, 12 (July), pp. 780-803. Reprinted in Russell (1917), ch. 1.
- RUSSELL, Bertrand (1914b) *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*, Chicago, Open Court Publishing Co. Republished by Routledge in 1993, with an introduction by John G. Slater.
- RUSSELL, Bertrand (1917) *Mysticism and Logic, and Other Essays*, New York: Longmans, Green, and Co. Ch. 1 reprints Russell (1914a).
- RUSSELL, Bertrand (1918) *The Philosophy of Logical Atomism*, London: Allen and Unwin. First published in *The Monist*, 28 (Oct. 1918), pp. 495-527; 29 (Jan., April, July 1919), pp. 32-63, 190-222, 345-80. Reprinted in Russell (1986), ch. 17.
- RUSSELL, Bertrand (1919a) *Introduction to Mathematical Philosophy*, London: Allen and Unwin. Second edition, 1920.
- RUSSELL, Bertrand (1919b) "On Propositions: What They Are and How They Mean", *Aristotelian Society Supplementary Volume*, 2, pp. 1-43. Reprinted in Russell (1986), ch. 20, pp. 276-306.
- RUSSELL, Bertrand (1921) *The Analysis of Mind*, London, Allen and Unwin. German translation by Kurt Grelling, *Die Analyse des Geistes*, Leipzig: Meiner Verlag: 1927.
- RUSSELL, Bertrand (1924) *Icarus; or, The Future of Science*, in To-day and To-morrow series, New York: E.P. Dutton and Company. Translated into German by Franz Arens as *Ikarus oder die Zukunft der Wissenschaft*, Munich: Drei Masken Verlag, 1926.
- RUSSELL, Bertrand (1925) *The ABC of Relativity*, London: George Allen and Unwin. Translated into German by Kurt Grelling as *Das ABC der Relativitätstheorie*, Munich: Drei Masken Verlag, 1928.
- RUSSELL, Bertrand (1927a) *The Analysis of Matter*, International Library of Psychology, Philosophy and Scientific Method, New York/London: Harcourt K. Paul, Trench, Trubner. Republished 2007 Nottingham: Spokesman. Translated into German by Kurt Grelling as *Philosophie der Materie*, Leipzig: B.G. Teubner, 1929.

- RUSSELL, Bertrand (1927b) *An Outline of Philosophy*, London: G. Allen and Unwin. Published in USA as *Philosophy*, New York: W.W. Norton, 1927; translated into German by Kurt Grelling as *Mensch und Welt: Grundriss der Philosophie*, Munich: Drei Masken Verlag, 1930.
- RUSSELL, Bertrand (1928) *Sceptical Essays*, Boston/London: George Allen and Unwin. Translated into German by Karl Wolfskehl as *Wissen und Wahn: Skeptische Essays*, Munich: Drei Masken Verlag, 1930.
- RUSSELL, Bertrand (1936) "The Limits of Empiricism", *Proceedings of the Aristotelian Society*, 34, pp. 131-150. Reprinted in Russell (1996), ch. 42.
- RUSSELL, Bertrand (1940) *An Inquiry into Meaning and Truth*, The William James Lectures, 1940, London: Allen and Unwin.
- RUSSELL, Bertrand (1959) *My Philosophical Development*, London: Allen and Unwin.
- RUSSELL, Bertrand (1967-1969) *Autobiography*, London and Boston/New York, George Allen and Unwin/Simon and Schuster. 3 Vols., London: George Allen and Unwin/Boston: Little Brown and Co (vols. 1, 2). New York: Simon and Schuster (vol. 3). Reprinted in Routledge Classics 2009.
- RUSSELL, Bertrand (1973) *Essays in Analysis*, London: Allen und Unwin. ed. by Douglas Lackey.
- RUSSELL, Bertrand (1986) *The Philosophy of Logical Atomism and Other Essays, 1914-1919*, ed. by John G. Slater, volume 8 of *The Collected Papers of Bertrand Russell*, London: Allen and Unwin.
- RUSSELL, Bertrand (1992) *Logical and Philosophical Papers, 1909-1913*, ed. by John G. Slater, volume 6 of *The Collected Papers of Bertrand Russell*, London: Routledge.
- RUSSELL, Bertrand (1994) *Foundations of logic, 1903-1905*, ed. by Alasdair Urquhart vol. 4 of *The Collected Papers of Bertrand Russell*, London: Routledge.
- RUSSELL, Bertrand (1996) *A Fresh Look at Empiricism, 1927-1942*, ed. by J. G. Slater and P. Köllner, vol. 10 of *The Collected Papers of Bertrand Russell*, London, New York: Routledge.
- RUSSELL, Bertrand (2014) *Toward "Principia Mathematica", 1905-1908*, ed. by Gregory H. vol. 5 of *The Collected Papers of Bertrand Russell*, London: Routledge.
- SCHILPP, Paul Arthur, ed., (1944) *The Philosophy of Bertrand Russell*, *The Library of Living Philosophers*, Evanston; Chicago: Northwestern University.
- SCHLICK, Moritz (1910) "Das Wesen der Wahrheit nach der modernen Logik", *Vierteljahrsschrift für wissenschaftliche Philosophie und Soziologie*, 34, pp. 386-377. Translated as "The Nature of Truth of Modern Logic", in Schlick (1979), ch. 3, pp. 41-103.
- SCHLICK, Moritz (1925) *Allgemeine Erkenntnislehre*, Berlin: Springer. Second edition, reprinted as Schlick, Moritz with Stadler, Friedrich and Wendel, Hans-Jürgen (series eds.) (2009), *Moritz Schlick Gesamtausgabe Abteilung I, Band 1: Allgemeine Erkenntnislehre*, Vienna: Springer; edited by Hans Jürgen Wendel and Fynn Ole Engler. English translation by Albert E. Blumberg as *General Theory of Knowledge*, La Salle, IL: Open Court, 1985.
- SCHLICK, Moritz (1927) "[Rezension von:] Bertrand Russell, *Die Probleme der Philosophie*, translated by Paul Hertz, (Weltkreis-Verlag, 1926)", in *Die Naturwissenschaften*, vol. 15, p. 626. Reprinted in Schlick (2008), pp. 81-4.
- SCHLICK, Moritz (1979) *Philosophical Papers, Vol. I (1909-1922)*, Dordrecht/Boston: D. Reidel. Edited by B.F. van de Velde-Schlick and H.L. Mulder.
- SCHLICK, Moritz (2008), with Stadler, Friedrich and Wendel, Hans-Jürgen (series eds.), *Moritz Schlick, Kritische Gesamtausgabe Abteilung I, Band 6: Die Wiener Zeit, Aufsätze, Beiträge, Rezensionen 1926-1936*, Vienna/New York: Springer. Edited by Johannes Friedl and Heiner Rutte.

- SCHÜTTE, Kurt (1965a) "Eine Grenze für die Beweisbarkeit der transfiniten Induktion in der verzweigten Typenlogik", *Archiv für mathematische Logik und Grundlagenforschung*, vol. 7, pp. 45-60.
- SCHÜTTE, Kurt (1965b) "Predicative Well-Orderings", in Crossley John N. and Dummett, Michael A.E., eds., *Formal Systems and Recursive Functions*, Amsterdam: North-Holland, p. 280-303.
- SHEFFER, Henry Maurice (1908) *A Program of Philosophy, Based on Modern Logic*, Harvard Dissertation.
- SHIEH, Sanford (2000) "Review of Hao Wang, *A Logical Journey: From Gödel to Philosophy*", *Erkenntnis*, 52, pp. 109-115.
- SHIEH, Sanford (2013) "Modality", in Beaney, ed., (2013) pp. 1042-1081.
- SHIEH, Sanford, manuscript, *Modality in Early Analytic Philosophy*.
- STADLER, Friedrich (2001) *The Vienna Circle. Studies in the Origins, Development, and Influence of Logical Empiricism*, Vienna, Springer. Translation of *Studien zum Wiener Kreis. Ursprung, Entwicklung und Wirkung des Logischen Empirismus im Kontext*, Frankfurt am Main: Suhrkamp, 1997.
- STEVENS, Graham (2005) *The Russellian Origins of Analytical Philosophy: Bertrand Russell and the Unity of the Proposition*, Routledge Studies in Twentieth-Century Philosophy, London/New York: Routledge.
- STEVENS, Graham (2008) "Russell and the Unity of the Proposition", *Philosophy Compass*, vol. 3(3), pp. 491-506.
- STUBENBERG, Leopold (2014) "Neutral Monism", in *The Stanford Encyclopedia of Philosophy*, E. Zalta, ed., at <http://plato.stanford.edu/entries/neutral-monism/>.
- SULLIVAN, Peter and JOHNSTON, Colin (forthcoming) "Judgements, Facts, and Propositions: Theories of Truth in Russell, Wittgenstein and Ramsey", in Glanzberg, Michael, ed., *The Oxford Handbook of Truth*, Oxford: Oxford University Press.
- TARSKI, Alfred (1931) "Sur les ensembles définissables de nombres réels", *Fundamenta Mathematicae*, 17, pp. 210-239.
- TARSKI, Alfred (1935) "Der Wahrheitsbegriff in den formalisierten Sprachen", *Studia Philosophica*, 1, pp. 261-405. German translation, with a postscript, of a 1933 paper in Polish. English translation by J.H. Woodger, "The Concept of Truth in Formalized Languages", in Corcoran, John, ed., *Logic, Semantics, Metamathematics. Papers from 1923 to 1938*, Indianapolis: Hackett, pp. 152-278.
- TAUSSKY-TODD, Olga (1987) "Remembrances of Kurt Gödel", in Weingartner, Paul and Schmetterer, Leopold, eds., *Gödel Remembered*, Napoli, Bibliopolis, pp. 31-41.
- TOLEDO, Sue (2011) "Sue Toledo's notes of her conversations with Gödel in 1972-1975", in Kennedy, Juliette and Kossak, Roman eds., (2011) pp. 200-207. Notes ed. by Juliette Kennedy.
- TRAVIS, Charles (2013) *Perception: Essays after Frege*, Oxford: Oxford University Press.
- VAN ATTEN, Mark (2014) "Gödel and intuitionism", in Dubucs, Jacques Paul and Bourdeau, Michel, eds., *Constructivity and Computability in Historical and Philosophical Perspective*, Dordrecht: Springer, pp. 169-214.
- VAN ATTEN, Mark and KENNEDY, Juliette, (2003) "On the philosophical development of Kurt Gödel", *The Bulletin of Symbolic Logic*, 9, pp. 425-476.
- VAN HEIJENOORT, Jan, ed., (1967/2000) *From Frege to Gödel: A Sourcebook in Mathematical Logic 1879-1931*, Cambridge, Mass.: Harvard University Press.

- VON NEUMANN, John (1929) “Über eine Widerspruchsfreiheitsfrage in der axiomatischen Mengenlehre”, in *Journal für die reine und angewandte Mathematik*, 160, pp. 227-241.
- WANG, Hao (1981) *Popular Lectures on Mathematical Logic*, New York: Van Nostrand Reinhold.
- WANG, Hao (1986) *Beyond Analytic Philosophy: Doing Justice to What We Know*, Cambridge, Mass.: MIT Press.
- WANG, Hao (1987) *Reflections of Kurt Gödel*, Cambridge, Mass.: MIT Press.
- WANG, Hao (1996) *A Logical Journey: From Gödel to Philosophy*, Cambridge, Mass.: MIT Press.
- WHITEHEAD, Alfred North and BERTRAND, Russell (1927) *Principia Mathematica*, Cambridge: Cambridge University Press, second edition; first edition 1910-1913.
- WITTGENSTEIN, Ludwig (1921) *Tractatus Logico-Philosophicus*, International library of psychology, philosophy, and scientific method, London, New York: Routledge and Kegan Paul. English translation by C.K. Ogden; reprinted with corrections 1933, 1993. Originally appeared in German in *Annalen der Naturphilosophie*, ed. by Wilhelm Ostwald, vol. 14, 1, 1921, Leipzig: Unesma.
- WITTGENSTEIN, Ludwig (1974) *On Certainty*, New York: Harper. English and German, translation of *Über Gewissheit*, ed. by G.E.M. Anscombe and G.H. von Wright, translated by Denis Paul and G.E.M. Anscombe.
- WITTGENSTEIN Ludwig (1978) *Remarks on the Foundations of Mathematics*, Cambridge, MA: MIT Press. Revised from the 1956 edition. Ed. by G.H. von Wright, R. Rhees, G.E.M. Anscombe, translated by G.E.M. Anscombe.
- WITTGENSTEIN Ludwig (2005) *Gesamtbriefwechsel = complete correspondence*. Electronic edition Innsbruck/Bergen/Oxford: Bergen Archives and Oxford University Press. Edited by Monica Seekircher, Brian McGuinness, Anton Unterkicher, Allan Janik and Walter Menthlagl.
- WITTGENSTEIN, Ludwig (2009) *Philosophische Untersuchungen = Philosophical Investigations*, Chichester, West Sussex, U.K. Malden, MA: Wiley-Blackwell, rev 4th edition. Edited and translated into English by G.E.M. Anscombe, P.M.S. Hacker, and Joachim Schulte.
- ZALABARDO José L., (2012) *Wittgenstein's Early Philosophy*, Oxford: Oxford University Press.
- ZERMELO, Ernst (1908) “Untersuchungen über die Grundlagen der Mengenlehre I”, *Mathematische Annalen*, 65, pp. 261-281. Reprinted and translated in Zermelo (2010), pp. 188-229.
- ZERMELO, Ernst (1930) “Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre”, *Fundamenta Mathematicae*, 16, pp. 29-47. Reprinted and translated in Zermelo (2010), pp. 400-431.
- ZERMELO, Ernst (2010) *Collected Works, Volume I*, ed. by Heinz-Dieter Ebbinghaus and Akihiro Kanamori, Berlin: Springer.