Joshua Schwartz
Dissertation Summary: "Quine, Disquotation, and Truth"
Committee: Michael Kremer (Chair), Josef Stern, and Kevin Davey.

In recent decades, an increasing number of analytic philosophers have been moved to write explicitly about the concept of truth in defense of a point of view labeled "Deflationism". At present, there are more than a handful of proposals for working out the deflationistic point of view. What unites Deflationists of different varieties is the view that the "philosophical" or "traditional" problem of truth — to define truth's nature — is a *Scheinproblem*; that there really is no such problem. My interest in this dissertation is not Deflationism, but the practice among Deflationists of claiming W. V. Quine as a philosophical progenitor. I demonstrate that this practice has bequeathed distorted representations of Quine's alethiology and that the recovery of Quine's views opens up neglected lines of philosophical inquiry.

Deflationistic readings of Quine are typically grounded in Quine's discussion of the "disquotational" feature of truth in "Truth and semantic ascent", which is a short section from his 1970 book *Philosophy of Logic*. In this section, Quine introduces the term "disquotation" to describe a feature of the truth predicate — the expression "is true" — that stands out in the following sentence:

(1) "Snow is white" is true if and only if snow is white.

This sentence is the paradigm of what philosophers call a "T-sentence". Its illustrative use in analytic philosophy derives from Alfred Tarski's hugely influential essay, "The Concept of Truth in Formalized Languages." About Tarski's paradigmatic T-sentence, Quine writes in *Philosophy of Logic*:

> By calling the sentence ["Snow is white"] true, we call snow white. The truth predicate is a device of disquotation.[1]

In later writings, Quine condenses this claim about the disquotational feature of the truth predicate to a shorter, more provocative saying: "Truth is disquotation".[2]

---

1. W. V. Quine, *Philosophy of Logic* (New Jersey: Prentice-Hall, 1970), 12.
2. W. V. Quine, *Quiddities: An Intermittently Philosophical Dictionary* (Cambridge, MA: Belknap Press

According to deflationistic presentations of Quine, the saying "Truth is disquotation" expresses Quine's view that a good account of the disquotational feature of the *truth predicate* exhausts what there is to say about *truth*. And, the story continues, Quine thinks that a good account of disquotation is an axiomatic theory of truth that looks like the collection of T-sentences patterned after (1); one T-sentence for each sentence of the language. What makes this approach to reading Quine deflationistic is, first, the assumption that a collection of T-sentences does not provide an answer to the question "What is truth?", and, second, the implication that no formulation correctly identifying truth's nature is forthcoming.

My dissertation challenges this widespread reading of Quine as proto-deflationist by arguing that, for Quine, the philosophical upshot of recognizing disquotation is *not* Deflationism, but, ultimately, a particular kind of inquiry into disquotation itself. In other words, I argue that deflationistic interpretations of Quine turn topsy-turvy his view of disquotation's significance: Quine does not think of disquotation or T-sentences as the *explanans* of a theory of truth, but, building on Tarski's work, he treats them as the *explanandum* of a rigorous inquiry into the nature of disquotation. Quine sets out this inquiry in an essay from around 1970 that philosophers have largely overlooked: "Truth and Disquotation."[3] In addition to making technical demands on readers, Quine's essay leaves us searching for the philosophical point of his investigation or the philosophical significance of the technical results that he reports. My dissertation reconstructs Quine's project in "Truth and Disquotation" and defends the view that his project aims to understand the ontology of disquotation. My conclusion is that, for Quine, the phenomenon of disquotation is entangled with mathematical ontology. I end the dissertation by representing Quine's alethiology and contemporary Deflationism as two ways of responding philosophically to the episode of conceptual change initiated by Tarski's mathematization of the concept of truth.

of Harvard University Press, 1987), "Truth", p. 213; W. V. Quine, *Pursuit of Truth*, rev. ed. (Cambridge, MA: Harvard University Press, 1992), §33, p. 80; W. V. Quine, "Reactions," in *Quine in Dialogue*, ed. D. Føllesdal and D. B. Quine (Cambridge, MA: Harvard University Press, 2008), 243.

3. W. V. Quine, "Truth and Disquotation," in *The Ways of Paradox*, revised and enlarged edition (Cambridge, MA: Harvard University Press, 1976).

To make sense of Quine's project in "Truth and Disquotation," we need to unearth the implicit Quinian framework that supports it. This excavation requires, as a first step, distinguishing Quine's concerns from those of contemporary Deflationists. My point of departure in the dissertation is a certain difficulty in the account of truth and disquotation that Quine presents in "Truth and semantic ascent". The difficulty stems from two key claims about the truth predicate that Quine advances in this text, which I call *Disquotation* and *Utility*. *Disquotation* is the idea that the truth predicate is "a device of disquotation". *Utility*, or Quine's utility claim, states that the use of generalized truth predications constitutes the utility of the truth predicate.[4] As Marian David points out in a paper on "Truth and semantic ascent", Quine's argument in this section suggests that *Disquotation* and *Utility* are closely related, but he fails to explicitly say how the two claims fit together.[5] The result of this omission is that Quine's text is ambiguous.

In Chapter 1, I distinguish three ways of understanding the relationship between *Disquotation* and *Utility*: (i) what I call the "standard interpretation" of Quine; (ii) Marian David's reading; and (iii) an approach that stems from Gary Ebbs.[6] The standard reading and David's interpretation understand Quine to be founding a deflationistic view of truth on *Disquotation* and *Utility*. Both readings, I argue, share a core assumption about Quine's alethiology: that, for Quine, the disquotational feature of the truth predicate *explains* the utility of generalized truth predications. What distinguishes (i) and (ii) is the kind of explanation that each attributes to Quine. According to the standard reading, disquotation provides something like a *semantical* explanation of generalized truth predications, whereas, on David's telling, Quine envisions a special kind of *syntactical* explanation, which David calls "the Ladder". I show that the problem with these two ways of reading Quine — David's and the standard one — is

---

4. Generalized truth predications are sentences that result from joining together a generalized noun phrase and the truth predicate. In *Philosophy of Logic*, Quine's example of a generalized truth predication is the following sentence: "Every sentence of the form '*p* or not *p*' is true." A more commonplace example of a generalized truth predication is: "Everything she said is true."

5. M. David, "Quine's Ladder: Two and a Half Pages from the *Philosophy of Logic*," *Midwest Studies in Philosophy* XXXII (2008): §1, pp. 276–277; §3, pp. 283–284.

6. Gary Ebbs, *Truth and Words* (Oxford: Oxford University Press, 2009).

that their common core assumption is false: both interpretations saddle Quine's idea of disquotation with explanatory ambitions that reach beyond anything that Quine explicitly says in "Truth and semantic ascent" or elsewhere.

My interpretation of "Truth and semantic ascent" builds on (iii), viz., Ebbs's approach to Quine's utility claim. According to my proposal, Quine's main idea in this section is that generalized truth predications "generalize on sentential position" in virtue of the existence of a "canonical deduction" from the generalized truth predication to its instances. The canonical deduction in question presupposes the availability of T-sentences as auxiliary premises. I argue that this technical role of disquotation in generalization-on does not imply that the disquotational feature of truth predicate explains generalized truth predications in the sense imagined by David and the standard reading. Therefore, the Ebbs-inspired approach I take up avoids the mistaken, core assumption of the two deflationistic interpretations.

Deflationistic readings of Quine assume that his comments about the disquotational feature of truth predicate and its utility are intended to "deflate" the traditional problem of identifying truth's nature. Distinguishing Quine's commitments from those of contemporary Deflationists allows me to question this assumption and to ask how Quine understands the problem of truth in the first place. My answer, articulated in Chapter 3 and Chapter 4, is that Quine identifies the problem of truth with the double-task of *explicating* disquotation and *understanding* it.

As I understand Quine, the use of an expression typically calls for explication as the result of linguistic pressures exerted by the development of science. In the first part of Chapter 3, taking a cue from Quine's discussion of truth in "On Austin's Method," I articulate this idea within the framework of Quine's *philosophical naturalism*. On the one hand, a scientist finds herself deploying an expression for a particular theoretical purpose, which is to say that the expression proves to have *utility for scientific theorizing*. On the other hand, insofar as the useful expression's home is ordinary language (or another specialized discourse), it might have additional uses that are needlessly defective from the point of view of the scientific enterprise in question.

Drawing on Quine's discussion of the ordered pair concept in §53 of *Word and Object*, I argue that, according to Quine, this linguistic tension gives rise to an *explicandum*: an expression whose primitive use works for the smooth development of scientific theory in some ways but against it in others. The Quinian response to this situation is *explication*: "finding a way of accomplishing those same purposes through other channels, using other and less troublesome forms of expression."[7] The new form of expression, adopted in place of the explicandum in practice or in theory, is called the *explicans*.

Quine thinks that the disquotational use of the truth predicate calls for explication in virtue of the utility of generalized truth predications for logical theory coupled with the threat of semantical paradox. On the one hand, the Quinian logician deploys the truth predicate to identify the truths of her subject, i.e., the logical truths, and she accomplishes this by generalizing on sentential position and quantifying over notational forms with respect to truth. This is the real meaning of Quine's utility claim: the use of generalized truth predications *in logic* constitutes the utility of the truth predicate. On the other hand, this specialized usage of the truth predicate can lead to paradox because of its roots in ordinary truth talk, which allows for the unrestricted construction of impredicative truth predications. The claim that I defend in Chapter 3 is that this situation fits Quine's description of what is paradigmatic about philosophical problems; the metalinguistic application of the truth predicate in logic — together with its singular, disquotational application — calls for explication.

Quine holds that Tarski's method of defining truth for the sentences of a particular formalized language, first set out in "The Concept of Truth in Formalized Languages," furnishes the desired explication of disquotation. This is one point of contact between the Quinian problem of truth and Tarski's work. However, the relationship between Tarski's work on truth and Quine's alethiology runs deeper than Tarski's provision of an explicans. If some term or form of expression becomes an explicandum at some point during the development of scientific inquiry, it does not follow that there is an additional philosophical task of "understanding" associated with the explicandum.

---

7. W. V. Quine, *Word and Object* (1960), §53, p. 260.

Why does Quine think that there is this kind of additional task in the case of disquotation? My answer, defended in the second part of Chapter 3, is that Quine finds the phenomenon of disquotation to be intrinsically puzzling, that is, puzzling apart from issues concerning the utility of a disquotational truth predicate for logic. I argue that this puzzle is rooted in Quine's view of Tarski's work on truth, in particular, the role of T-sentences in Tarski's argument in §§1–3 of "The Concept of Truth in Formalized Languages."

In Chapter 2, I contend that Tarski deploys T-sentences in a dual-fashion: constructively and destructively. On the one hand, in §1, Tarski describes T-sentences as "partial definitions of the truth of a sentence".[8] In this role — as "partial definitions" — Tarski's T-sentences initiate the task of giving a correct "semantical definition of truth" for the sentences of colloquial language. On the other hand, Tarski argues in §1 that this task bogs down *at its starting point*, that is, with the T-sentences themselves. The problem according to Tarski is that colloquial language, by dint of its "universality", allows the formulation of T-sentences that are *not* partial definitions; these "T-sentences", as he demonstrates, generate semantical paradox. This is the destructive role of Tarski's T-sentences. In §3 of "The Concept of Truth in Formalized Languages," Tarski extends the constructive role of T-sentences as partial definitions in order to justify his well-known proposal of a condition of "material adequacy" for an explicit definition of the true sentences of a predetermined formalized language. This condition is Convention **T**. In Chapter 2, I argue that Tarski's adoption of Convention **T** rests on the power of disquotation to uniquely determine the extension of the colloquial truth predicate.

This extension-determining power of the disquotational feature of the colloquial truth predicate makes disquotation maximally clear according to Quine. However, the destructive role of T-sentences in generating semantical paradox makes disquotation maximally unclear. The Quinian puzzle concerning disquotation is how disquotation

---

8. A. Tarski, "The Concept of Truth in Formalized Languages," in *Logic*, *Semantics*, *Metamathematics*, 2nd. ed., ed. J. Corcoran, trans. J. H. Woodger (Indianapolis: Hackett Publishing Company, 1983), §1, p. 155.

can make one and the same term maximally clear and maximally unclear.[9] I support this interpretation in Chapter 3 by considering passages from Quine's corpus that describe the clarity and unclarity of disquotation as aspects of single concrete phenomenon. The puzzle of disquotation, on my telling, underlies Quine's view that disquotation call for philosophical understanding.

But what kind of understanding of disquotation does Quine seek? To answer this question, I turn in Chapter 4 to Quine's project in "Truth and Disquotation." In this paper, Quine treats Tarski's explication of disquotation as a "going concern" that is subject to possible simplification. As Peter Hylton discusses at length in his book on Quine's philosophy, a good portion of Quine's philosophical writings aim to produce the most systematic, simple, and clear formulation of total science by showing how to formulate that (putative) knowledge within a *canonical notation*. Hylton describes this project as "metaphysics naturalized".[10] According to my reading, Quine's investigation of disquotation in "Truth and Disquotation" is a chapter of naturalized metaphysics in Hylton's sense.

In *Word and Object* and other familiar writings, Quine's canonical notation is the notation of first-order, quantification theory, and the logical framework of total science is classical, first-order logic. Tarski's method of truth-definition shows how to define a disquotational truth predicate for sentences expressed in Quine's canonical notation (in a metalanguage whose logical form is also canonical). The Tarskian technique is indirect: a (formal) truth predicate is introduced by an explicit definition in terms of a more general semantical relation — satisfaction. In the first section of "Truth and Disquotation," Quine claims that there is a pair of "complications" that attaches to Tarski's original definition of truth in terms of satisfaction: (i) Tarski's definition appeals to a collection of sequences of objects (from the domain of the object-language) of arbitrary finite length; (ii) the clause of Tarski's definition of satisfaction that pertains to formulas whose main logical operator is a quantifier-variable prefix is, Quine

---

9. My formulation of this perplexity is due to Michael Kremer.
10. P. Hylton, *Quine* (New York: Routledge, 2007), Chapter 9, p. 232.

thinks, "more devious and complex" than other clauses of the definition.[11] Since these two complications are occasioned by the quantificational structure of Tarski's object-language, Quine asks in "Truth and Disquotation" whether adapting Tarski's technique to *variable-free* formalisms eliminates (i) and (ii). In the main sections of his essay, Quine pursues this question in connection with two forms of variable-free notation: Schönfinkel's combinatory logic and Quine's own variable-free formulation of first-order logic, which he calls "predicate-functor logic".

In order to make philosophical sense of Quine's technical question, I put a question to Quine in turn: Why are the appeal to sequences and the allegedly unnatural clause for quantification "complications" worth thinking about? My answer is that, for Quine, (i) and (ii) are complications from the point of view of an explanation of disquotation. I argue that implicit in Quine's discussion is a distinction between two forms of disquotation, which I call "weak" and "strong" disquotation respectively. Weak and strong disquotation are properties of an inductive definition of truth that is materially adequate in Tarski's sense. Roughly speaking, a Tarski-style truth-definition is *strongly disquotational* if the definition alone completely captures the disquotational feature of the truth predicate. Otherwise, the truth-definition is *weakly disquotational*. In Chapter 4, I demonstrate that owing to the two complications identified by Quine, Tarski's definition of truth via satisfaction is weakly disquotational; the truth-definition captures disquotation only in conjunction with a theory of finite sequences.

Equipped with the distinction between weak and strong disquotation, I recast Quine's project: the question driving "Truth and Disquotation" is whether we can find a recursive characterization of truth that is strongly disquotational by adapting Tarski's method to variable-free "styles of logic" like combinatory logic and predicate-functor logic. In other words, I maintain that Quine wants to know whether weak disquotation is merely an artifact of quantifier-variable notation. This question makes sense, however, only if Quine thinks that quantificational notation itself is dispensable

---

11. Quine, "Truth and Disquotation," §I, p. 309.

in favor of some more basic formalism. Since quantifier-variable notation constitutes Quine's canonical notation, it might seem that my reading of "Truth and Disquotation" is mistaken. I respond to this concern by pointing out that Quine endorses a moderate pluralism of canonical notations in *Word and Object*: a logical formalism qualifies as canonical if it is intertranslatable with first-order, quantifier-variable notation.[12]

In Chapter 3, I argue that Quine's invention of predicate-functor logic constitutes a self-standing, homegrown example of Quinian explication. The explicandum in this case is the bound variable; the explicans is a small collection of operators on predicates — the predicate functors. Schönfinkel's combinatory logic also provides analysis of the bound variable. However, Quine maintains that Schönfinkel's system of combinators overproduces as an analysis of the bound variables of first-order logic because, according to Quine's view, the (untyped) combinators presuppose a mathematical ontology equivalent to "higher set theory", that is, the mathematical theory of infinite sets. By contrast, Quine holds that his own predicate-functor scheme is precisely adequate to the work of the bound variable. Quine's view of the differences between Schönfinkel's combinatory logic and predicate-functor logic rests on the fact that the latter is intertranslatable with standard, quantificational first-order languages, whereas the untyped combinators cannot be defined in terms of elementary logic alone.

Insofar as Quine holds that predicate functors reduce quantifier-variable notation to more basic terms, I propose in Chapter 4 that we should think of predicate-functor logic itself as a more austere canonical notation for science than classical quantification theory. Leaning on this proposal, I argue that there is a point of view available to Quine from which the use of quantificational notation is optional or "merely notational", namely, the point of view provided by treating predicate-functor logic as canonical. That is, predicate-functor logic furnishes Quine with a tool for probing the distinction between weak and strong disquotation. If it were possible to extend Tarski's technique to the predicate-functor setting and to obtain a form of strong

---

12. Quine, *Word and Object*, §39, p. 188.

disquotation that overcomes the "complications" attaching to Tarski's original truth-definition, then Quine would have a reason to characterize weak disquotation as an artifact of quantificational notation. Furthermore, if he could construct a strongly disquotational truth-definition for predicate-functor logic, then Quine would be able to eliminate the ontological assumption of finite sequences from the explanation of disquotation.

However, in §V of "Truth and Disquotation," Quine argues that a strongly disquotational truth-definition for predicate-functor logic is not possible. Drawing on unpublished writings and letters, I show that Quine thought at one point that he *had* obtained the desired simplification of Tarski's original truth-definition. I argue that the impetus for "Truth and Disquotation" was the realization that his "truth-definition" for predicate-functor logic was fundamentally flawed. Given this negative result, I contend that there is no need for Quine to revert to a predicate-functor style of canonical notation in order to explain disquotation; the best theory of disquotation is provided by Tarski's original, weakly disquotational truth-definition together with its ontology of finite sequences and its unnatural clause for quantification. This situation, I maintain, provides one reason for ascribing to Quine the view that disquotation is entangled with mathematical ontology.

A second reason concerns the results of Quine's investigation of disquotation in the setting of combinatory logic. In §II of "Truth and Disquotation," Quine reports a positive result: he characterizes a function, which he calls "designatum", and he shows that this characterization is a materially adequate definition of truth for the sentences of the language of combinatory logic. I show that Quine's definition of designatum overcomes the two "complications" attaching to Tarski's original truth-definition; Quine's definition of truth for combinatory logic does not explicitly appeal to a domain of finite sequences and it contains no unnatural clauses. Furthermore, I demonstrate that Quine's simplified truth-definition for combinatory logic exhibits strong disquotation. According to Quine's understanding, Schönfinkel's combinatory logic assumes a substantial mathematical ontology that is "equivalent to higher set

theory". If Quine's interpretation of the (untyped) combinators is correct, then it follows that the ontological commitments of his strongly disquotational truth-definition for combinatory logic exceed the modest theory of finite sequences required in Tarski's original quantificational setting. In other words, Quine thinks that if we strengthen the mathematical assumptions of the theory of disquotation, then we can upgrade from weak to strong disquotation. In Chapter 4, I present this view of Quine's as a second reason for ascribing to him the thesis that the phenomenon of disquotation and mathematical ontology are knotted together.

In my conclusion, I return to the differences between Quine's alethiology, which is centered on the problem of truth *as disquotation*, and contemporary Deflationism, whose primary target is the traditional problem of identifying truth's nature. Drawing on Leon Horsten's description of "the Tarskian Turn" in alethiology, I argue that we can think of Quine's views and those of contemporary Deflationists as two different ways of responding philosophically to Tarski's mathematization of the concept of truth. I suggest that the project of contemporary Deflationism is best seen as an attempt to build on Tarski's work in order to provide a mathematical model (not in the sense of model theory) of the natural language truth predicate or the ordinary concept of truth. In contrast, I maintain that Quine's view of truth and disquotation looks inward to the episode of conceptual change itself, i.e., to Tarski's mathematization of truth. The Quinian path opened up by my dissertation leads us to questions about the nature of concept-formation in mathematical logic, in particular, whether episodes of conceptual change in logic illuminate the distinctive business of mathematics more generally.