TRUTH, RAMSIFICATION, AND THE PLURALIST’S REVENGE

Cory D. Wright

Functionalists about truth employ Ramsification to produce an implicit definition of the theoretical term *true*, but doing so requires determining that the theory introducing that term is itself true. A variety of putative dissolutions to this problem of epistemic circularity are shown to be unsatisfactory. One solution is offered on functionalists’ behalf, though it has the upshot that they must tread on their anti-pluralist commitments.

1. Introduction

Truth-theorists are sometimes asked what a theory of truth is about. In answering ‘truth’, we go around in a very small circle. A less loopy answer might instead begin with a little more philosophical spadework. For instance, we might try to gainfully reinterpret the question as a question about how the terms used to talk about truth—such as *true* or *true*—are to be defined. After all, for any given theory of truth $T$, such terms are theoretical terms; and questions about how to define theoretical terms have well-known answers.

Inflationists often take this route, claiming that *true* attributes a property which consists in being $F$. Functionalism about truth is one such inflationary theory, although its central thesis is the further claim that being $F$ is a functional kind rather than a structural one. More specifically,

$$(\text{fnct}) \text{ true refers to the single higher-order functional role property } F \text{ of having lower-order properties } p_1, \ldots, p_n \text{ that realize it.}$$

To arrive at this thesis, functionalists suppose that *true* can be defined implicitly via the quasi-formal technique of Ramsification.

---

1Given that philosophers of language don’t have a consistent and unified theoretical vocabulary for discussing semantic, semiotic, and alethic relations (reference, deixis, signification, designation, predication, denotation, profiling, satisfaction, etc.), I’ll generally use terms for reference as the most neutral genus-level family of terms allowing us to bypass thorny questions about which are the most grammatically appropriate to invoke and when. In certain contexts, attribution—although clearly not synonymous with terms for reference—will be preferable.

2Functionalism about truth, which sometimes goes by the misnomer alethic functionalism, has been developed primarily by Lynch [2000, 2004, 2005, 2009]. The theory was anticipated by Lafleur [1941], to a limited extent, and again in a more recognizable form by Petit [1996]. For additional exposition, see Devlin [2003], Wright [2005], and Sher [2005].

---
My contention is that theories of truth employing Ramsification face a problem of epistemic circularity. The problem arises because any such implicit definition proceeds, at least in part, on the basis of explicit decisions that certain sentences containing the *definiendum* are themselves true. Rather than simply dispensing with Ramsification itself [cf. Lynch 2009], I propose nine putative dissolutions to the problem. After explaining why each fails, I offer a positive solution on behalf of functionalists. The solution, however, requires of functionalists that they slough off certain monistic assumptions. Since both the problem and the solution generalize to any theory of truth relying on Ramsification to define the theoretical terms used to talk about truth, pluralism becomes a pressing issue. I leave it as an open question whether pluralists have a viable position themselves.

2. Why Implicit Definition?

Perhaps the most straightforward way to define theoretical terms is simply to provide explicit and noncircular definitions of the form

\[(\text{defn}) \quad x \ldots x \ldots =_{df} \ldots,\]

where \(\ldots x \ldots\) is some *definiendum* involving the unique term being defined, \(=_{df}\) is the relation of definitional equation, and \(\ldots\) is the *definiens* in terms of which \(x\) is explicitly defined but doesn’t occur.

Theoretical terms often resist explicit and noncircular definitions, however. Indeed, scientific theories are renowned for occasioning terms that have technical or specialized senses, or express unfamiliar or exotic conceptualizations, or stand for unobservable posits or abstracta of questionable repute: e.g., *caloric*, *dark matter*, or *superstring* in physical theory, *fitness* or *race* in biology, and *sense-datum*, *id*, or *intelligence* in the psychological sciences. Call these recalcitrant terms the \(\tau\)-terms of a scientific theory \(T\). A \(\tau\)-term-introducing theory \(T\) needn’t be exclusively scientific, however. Many of our more interesting philosophical concepts are those whose expression—e.g., *causation*, *representation*, *person*, and *well-being*—has also proven resistant to explicit noncircular definition.

The reasons for resistance vary widely. Often, they are purpose-relative. Those \(\tau\)-terms introduced stipulatively, for example, can simply fail to achieve their intended aim; those introduced descriptively may underestimate the wide variety of actual usages; those introduced ampliatively can fail to constitute a sufficient theoretical improvement; and so forth. In other cases, the reasons for resistance are term-specific. For example, slingshot arguments await those correspondence theorists attempting to define *truth* in terms of a structural relation of correspondence between truth-bearers and facts. For a term like *coherence*, the difficulty is due to a certain lack of mathematical precision among comparative approaches to the analysis of *coherence* [Millgram 2000: 82–3].
In still other cases, attempts at explicit non-circular definition are further encumbered because the τ-term targeted for definition has quotidian counterparts that enjoy common and versatile usage in ordinary discourse. Such cases seemingly induce a feeling of trying to define something that is paradoxically strange but familiar, profound yet mundane [Lynch 2005: 29; Næss 1938: 159–60]. One observes this in even the earliest and most elementary of the Platonic dialogues—Socrates’s attempt to formulate ampliative definitions of friendship in the Lysis, for example. And as the number of distinct counterparts increases, the encumbrance is further compounded. The difficulty in explicitly defining the τ-term sentence in linguistic theory, for example, is due not so much to its merely having additional senses, but to its having hundreds of them [Ries 1931].

In the case of truth, Davidson [1996] urged that the appropriate response is to quit: attempts to give any kind of general definition are unwittingly involved in folly simply because the term truth—as he took Tarski to have shown—is indefinable. For some, Davidson’s brand of primitivism seems unduly pessimistic. An alternative strategy for managing difficulties incurred from attempts at explicit and noncircular definitions of τ-terms is simply to offer implicit definitions instead. Rather than equating a definieund with some definitiens, implicit definitions take the form of certain true sentences σ₁, σ₂, . . . that are constitutive of T and in which the definiendum occurs. For example, τ-terms like point, line, and radius, which occur in the axioms of Euclidian geometry, are commonly said to be implicitly defined by those axioms. The assumption underlying this strategy is that, in reckoning that T is true, one assigns, implicitly, its τ-terms the meanings they would need to have in order for T to be true. Accordingly, it’s sufficient for determining what a τ-term means that the theorist determine some, perhaps even many, of the true sentences in which it features as a subsentential component.

The strategy of implicit definition forms the basis for the quasi-formal technique of Ramsification, which has long been offered as a way to define τ-terms by exchanging them—purportedly without loss of meaning—for a ‘street-level’ vocabulary. Since the meanings of those τ-terms weren’t well understood in the first place, then quantifying over the variables that replace them stands to produce a definitive sentence serving as a descriptively adequate substituend for the original theory.

3. Defining Alethic Terms Implicitly

3.1 Ramsification

The standard version of Ramsification is the one popularized by Lewis [1970],³ which begins by amassing whatever principles P₁, P₂, . . . are constitutive of the postulate of T or otherwise relevant to its introduction

³Modifications and alternatives to Lewis’s version abound; see, e.g., Craig [1953], Martin [1966], Bohnert [1967], Bedard [1993], Hawthorne [1994], and Horwich [1997]. None of them circumvents the problem of epistemic circularity discussed herein.
of the target \( \tau \)-term. Having fixed upon this set of principles, the second step is to order and then conjoin them to form a single sentence which is materially equivalent to \((T: P_1, \ldots, P_n)\). Lewis had us rewrite the result as

\[
(T) \quad R(\tau_1, \ldots, \tau_n, o_1, \ldots, o_n),
\]

so as to isolate the \( \tau \)-terms targeted for implicit definition. (For the purpose of homogenizing the variables that will appear in forthcoming derivations, Lewis also proposed that we convert all functors and predicative \( \tau \)-terms in the postulate of \( T \) to their corresponding nominalizations, i.e., \( \tau \)-name\(_1, \ldots, \tau \)-name\(_n\)).

Any such implicit definition necessitates a large register of additional terms, against which the meaning and use of \( \tau \)-terms can be situated. After all, the string \( \tau \)-name\(_1, \ldots, \tau \)-name\(_n\) would be quite useless otherwise—nearly incomprehensible, practically incommunicable, and certainly ungrammatical. The remainder of \( R \) therefore consists in what Lewis called old-or-original-or-other terms. Unfortunately, Lewis said little else about how to characterize these \( o \)-terms, though the very idea of an \( o \)-term is perhaps sufficiently intuitive. If need be, we can precisify it as follows. Let an expression \( x \) be an \( o \)-term in a given language \( \mathcal{L} \) iff \( x \) is a symbolic structure \( \Sigma \) with unit status in the grammar of \( \mathcal{L} \) prior to the formulation of \( T \), where \( [\Sigma] \) exhaustively consists in a phonological unit \([p]\) that (literally) calls to mind a semantic unit \([s]\) for speakers of \( \mathcal{L} \) (i.e. \([s] \rightarrow [p] \)). An \( o \)-sentence is then any sentence of \( \mathcal{L} \) free of \( \tau \)-terms.

The next step is to replace every \( \tau \)-name instance with a corresponding subscripted variable \((x_1, \ldots, x_n)\) ranging over individuals in the domain of \( R \). In the case at hand, all terms pertaining to truth, truths, truth-talk, judgments-of-truth, etc. are stripped out, as well as terms that are potentially interdefinable with them—assert, real, fact, etc. Doing so results in an open sentence,

\[
(T') \quad R(x_1, \ldots, x_n, o_1, \ldots, o_n),
\]

which Lewis called the realization formula of \( T \). Basically, \( T' \) is a specification of what must obtain for the postulate of \( T \) to be
realized—i.e., a ‘job description’ as it’s sometimes put. And the postulate of ℰ is realized just in case ℰ′ is satisfied by any ordered n-tuple of entities—i.e., some realizer ρi that plays the role or does the job identified and individuated by ℰ′. It is uniquely realized just in case exactly one ordered n-tuple satisfies ℰ′, multiply realized if more than one does, and unrealized if none does. Whether a theory is unrealized, uniquely realized, or multiply realized depends on what there is.

Articulating the sentence stating that ℰ is so realized—namely, its Ramsey sentence ℰR—is achieved by prefixing, for each subscripted variable xi ∈ ℰ, an existential quantifier binding it.

\[(T_R) \quad \exists x_1, \ldots, \exists x_n[R(x_1, \ldots, x_n, o_1, \ldots, o_n)].\]

The Ramsey sentence simply specifies that there exists some such realizer ρi. And the point of doing so, of course, is to position the theorist to be able to make the same ontic commitments that would otherwise be incurred by asserting or endorsing ℰ.5

3.2 Application to Theories of Truth

From ℰR, the conditions under which some ρi is possessed can then be given by embedding the Ramsey sentence in the requisite biconditional,

\[(T_L) \quad \rho_i(a) \equiv \exists x_1, \ldots, \exists x_n [R(x_1, \ldots, x_n, o_1, \ldots, o_n) & a \text{ has } x_1, \ldots, x_n].\]

ℰL states that an individual a has some ρi denoted by the (nominalized) τ-term when and only when the variable xj replacing it both is R-related to certain other terms in the postulate of ℰ and is had by a. For functionalists about truth, the individual is some truth-bearer σ and the corresponding biconditional states that σ has some alethic property ρi that realizes the F-role just in case the extant value of the variable xj replacing truth is both τ-named by the terms of the theory en masse and is had by σ:

\[(T_L^f) \quad \sigma \text{ has some alethic property } \rho_i \text{ realizing } F \\
\quad \equiv \exists x_1, \ldots, \exists x_n [R(x_1, \ldots, x_n, o_1, \ldots, o_n) & \sigma \text{ has } x_1, \ldots, x_n].\]

5Of course, the derivation from ℰ to ℰR yields a less informative and more abstract sentence (roughly, in the same sense that the sentence there is at least one truth logically entails the less informative and more abstract sentence there is at least one entity). Yet, as Ramsey famously noted, the magnitude of ℰR is no less powerful, given that it makes all the same predictions and inferential connections between observation sentences. Furthermore, as Bohnert [1967] less famously noted, the flight from informativeness is kept at a minimum. For instance, logically equivalent sentences, such as the conjunction of shorter conjunctions,

\[(T_R^w) \quad \exists x_1[R(x_1, o_1, \ldots, o_n)] & \ldots & \exists x_n[R(x_n, o_1, \ldots, o_n)],\]

are not necessarily as expressively powerful as ℰR; for τ-names may irrigidly denote where various realization formulae fall within the scope of different quantifiers. In the case of pluralism about truth, irrigid denotation might be something comfortably accommodated or tolerable.
From $T'_L$, functionalists can then make explicit the conditions under which $s$ is true:

$$ (T''_L) \quad \text{s is true} \equiv \text{s has some alethic property } \rho, \text{ realizing } F. $$

Yet, as Lynch [2004, 2005] rightly noted, neither $T'_L$ nor $T''_L$ entails a commitment about the ‘deep’ metaphysical nature of truth itself; these biconditionals only specify necessary and sufficient conditions. Given the choice, functionalists identify truth with the $F$-role itself—not with whatever ordered $n$-tuples happen to realize the property of being the $F$-role for a given sector of (truth-apt) discourse [Pettit 1996: 886]. More specifically,

$$ (T''_L) \quad \text{true: } s \text{ is true} =_{df} s \text{ has the property of having some realizer } \rho, \text{ of the } F\text{-role}. $$

Subsequently, the inferential route to the central thesis of functionalism about truth should be clearer. To say that the $\tau$-name $\text{truth}$ denotes a functional kind $F$ is to say that truth is the property of having a property that plays the $F$-role.

In turn, it should also be clearer why functionalism about truth proves to be a kind of inflationary theory to rival all others, including the correspondence theory. $T''_L$ entails (although is not entailed by) the inflationist claim that true refers to $F$ and is explicitly and non-circularly definable in terms of it,

$$ (\text{infl}) \quad \text{true: } s \text{ is true} =_{df} s = F, $$

given only the assumption that properties consisting in being $F$ just are, or are reducible to, $F$.

### 3.3 Monism versus Pluralism

The choice of identifying truth with the $F$-role in lieu of its realizers allows functionalists to bypass several problems thought to beset monistic and pluralistic theories. On one hand, monists claiming that ‘truth is one’ take the nature of truth to be uniform across all (truth-apt) sectors of discourse; yet, some candidate properties and relations are intuitively plausible for some sectors, but far less so in others. This is the so-called scope problem.\(^6\)

For example, a monistic theory proposing that the truth of all true sentences in cartographical discourse consists in being structurally homomorphic may be correct, fecund, provide elegant explanations, and so forth. Yet, homomorphism is less intuitively plausible as a candidate for what truth consists in for true sentences in mathematics, comedy, comedy, comedy, comedy, comedy, comedy, comedy, comedy, comedy, comedy, comedy, comedy, comedy, comedy.

---

\(^6\)For further discussion of the scope problem, see Sher [1998, 2004] and Lynch [2004]. For discussion of the problems faced by pluralists about truth, including equivocation, generalization, and mixed inference, see, e.g., Lynch [2000, 2004], Wright [2005], Sher [2005], and Pedersen [2006].
aesthetics and gastronomics, ethics, and a variety of other sectors of discourse.

On the other hand, consider pluralists who claim that ‘truth is many’. For instance, one might identify the “deep” metaphysical nature of truth itself with the lower-order realizers $r_1, r_2, \text{ and } r_3$ (e.g., accuracy, super-assertibility, and coherence) of $F$ in different discourse sectors. Yet claims that truth denotes not one thing but many putatively face problems of equivocation, generalization, and mixed inference owing to their proliferation of truth properties. For example, if the truth of psychological and moral sentences consist in different properties, then the inference from the premises Joe Louis is vain and vanity is a moral deficit to the conclusion Joe Louis has a moral deficit may fail to be a valid inference if no single uniform property of truth is being preserved.

Functionalism about truth attempts to circumvent these one–many problems. By its central thesis $fnct$, there is a single higher-order functional role property $F$ of having many lower-order properties $r_1, \ldots, r_n$ that realize it. Subsequently, functionalists take us a great distance in showing how we might try to accommodate both pluralist and monist views about truth while avoiding the problems of each. $T_0$ still allows for the possibility that $F$ itself is multiply realized—i.e., what Devlin [2003: 59] correctly calls polyalethism—in so far as what realizes the higher-order role property $F$ may differ across sectors of discourse. But it precludes the possibility that the postulate of $T$ is multiply realized, since the $F$-role is the single, unique property denoted by the singular $t$-name targeted for analysis. And Ramsification is the lynchpin in all of this.

4. Realization and Epistemic Circularity

Advocates of Ramsification often follow Lewis in supposing that $T$ is defective if $T'$ is unrealized or multiply realized. The sense of defective, here, is the sense in which the theory attempts to introduce a theoretical term to name something specific, but nothing is so named in the former case (i.e., the definiendum has $\emptyset$ as its extension) and the thing named in the latter case is we know not what (i.e., the definiendum suffers from referential indeterminacy). Consequently, only in cases of unique realization—whereby the

---

7Some of these ‘problems’ arise from a faulty characterization of the nature of semantic and conceptual ambiguity, and so are easily dissolved.

8The motivation to employ the functionalist apparatus of roles and realizers in the first place would be far less persuasive if the set of realizers turned out to be a set of cardinality one.

9Pace Lewis [1970: 432–3], the specific type of theoretical defect is falsity. Lewis’s claim is fairly straightforward in cases where $T$ is unrealized. The $t$-names miasma and phlogiston, for example, are introduced by the theories of miasmatic transmission and phlogistification under the assumption that the existential quantifier binding each $x_i \in R$ implies existence; but we count those theories as being false since they imply the existence of non-extant entities. Likewise, if $T$ is some version of monism or pluralism about truth and no ordered $n$-tuple of entities satisfies their respective realization formulae, then the sentence there’s either one or many ways of being true is false. Cases of $T$ being multiply realized are far less straightforward.

We can certainly concur that multiple realization is less desirable than unique realization, since being unable to non-arbitrarily (re-)identify the $i$th entity of the ordered $n$-tuple of entities satisfies their respective realization formulae, then the sentence there’s either one or many ways of being true is false. Cases of $T$ being multiply realized are far less straightforward.

10We can certainly concur that multiple realization is less desirable than unique realization, since being unable to non-arbitrarily (re-)identify the $i$th entity of the ordered $n$-tuple of the realization by the $i$th term of $T'$ is surely a defect. But it’s unclear why that defect is one of falsity. As Bedard [1993: 506] notes, extreme complexity and weakness—like multiple realization—is unattractive; yet, we don’t count $T$ as false simply because it’s not simple or strong enough. In any case, given that functionalism is a version of monism about truth (there is but one $F$-role per $T_0$; see [Wright 2005: 13–15]), let us momentarily bracket these complications.
elements of the ordered $n$-tuple are the precisely the ones $\tau$-named by their corresponding terms in $R$—can the initial postulate of $T$ can be fully recovered from the Ramsey sentence in conjunction with the Carnap sentence,

$$(T_C) \quad \exists x_1, \ldots, \exists x_n(x_1, \ldots, x_n, o_1, \ldots, o_n) \supset R(\tau_1, \ldots, \tau_n, o_1, \ldots, o_n).$$

Unique realization is thus the standard of correctness we should insist on, unless it turns out to be a standard too stringent to meet—that was Lewis’s insight.

For any given $T$, determining whether that standard is met requires determining, a fortiori, that the antecedent of the Carnap sentence with uniqueness,

$$(T'_C) \quad (\exists! y_1, \ldots, \exists! y_n)(\forall x_1, \ldots, \forall x_n) [R(x_1, \ldots, x_n, o_1, \ldots, o_n)] \equiv (y_1 = x_1 \land \cdots \land y_n = x_n) \supset R(\tau_1, \ldots, \tau_n, o_1, \ldots, o_n),$$

holds. (For our purposes, we can read $T'_C$ as stating that, if there exists a single unique $n$-tuple satisfying the functionally specified profile, then the theory $\tau$-names it.)

In turn, determining that the antecedent holds involves determining, inter alia, the truth-values of all conjuncts of $R$. After all, recall that one of the first steps in producing the Ramsey sentence is that the postulate of $T$ be rewritten as a materially equivalent conjunction $R$. Assuming classical logic, $T$ will be false if any given conjunct of $R$ is false. Where $T$ is false, so too will be the realization formula $T'_R$, the Ramsey sentence $T'_R$, and the Carnap sentence with uniqueness $T'_C$ derived from it. And where any of these is false, we have good reason to think that the definitendum fails to denote. With Ramsification, what you put in is what you get out.

And so it’s here that the problem of epistemic circularity rears. Functionalists about truth implicitly define their (nominalized) $\tau$-term $true$ by using Ramsification to produce a sentence that indicates its denotation. But any implicit definition proceeds on the basis of explicit decisions that the principles constitutive of $T$ are themselves true. Hence the circularity. In turn, making any explicit decisions that they are true requires already knowing in advance what truth is. Hence the epistemic circularity. If we suppose further that knowing in advance what truth is entails knowing what the $\tau$-name $true$ denotes, then it becomes unclear why functionalists about truth ever required an implicit definition via Ramsification in the first place. Hence the problem. Moreover, we might add, failure to determine the truth-values of the conjuncts of $R$ would position the theorist to be unable to ascertain whether the implicit definition can be maintained for any and all permutations of $(T: P_1, \ldots, P_n)$ or other intra-theoretic refinements.
5. Putative Dissolutions

It's surely no consolation to functionalists that the problem of epistemic circularity generalizes to any theory of truth invoking Ramsification to implicitly define true. Interestingly though, the problem is not a fatal one—either for the functionalist or any other truth-theorist. Before providing one solution, however, let me briefly canvass nine ways one might attempt to dissolve the problem, each of which is deemed unsatisfactory.

First, one might argue that there's no problem because theorists needn't do anything beyond observing that the Ramsified \( t \)-terms of a true theory are implicitly defined by its axioms. This putative dissolution misses the mark, and widely. Yes, the terms are so defined; but any truth-theorist resting their analysis on such an observation is one whose 'employment' of the technique is little more than a gratuitous logical drill. Truth-theorists employ Ramsification because they are interested in advancing their theories, not because they are interested in giving an exposition of the methods they could possibly use to do so. At the very least, functionalists about truth can hardly be accused of such daffidence. They employ Ramsification to show that the \( t \)-name truth they introduce denotes what they say it does, not to show that one might hypothetically define it as such were there to eventually be a true theory about truth.

Second, one might argue that truth-theorists need not determine the truth-values of the conjuncts of \( R \). Instead, they need only endorse or assert \( T \), plus any of the derivations produced as the output of Ramsifying over it. The problem with this putative dissolution is the insufficiency of all such asseveration; for mere endorsements or assertions of a theory, however vehement, are no guarantee against the possibility of its falsity. And theorists endorsing or asserting a false theory are theorists who fail to implicitly define a target \( t \)-term via Ramsification. Consequently, it won't do to argue that the problem of epistemic circularity dissolves once we distinguish the requirement (i) that all conjuncts of \( R \) must be true from the requirement (ii) that all conjuncts must be determined to be true.\(^\text{10}\) The distinction, while clear, fails to do the work demanded of it. It’s tantamount to claiming that the target \( t \)-term will be implicitly defined by merely articulating the constituent principles of \( T \), as long as requirement (i)—but not necessarily (ii)—is satisfied. Now plainly, the material point isn’t about the articulation of \( T \), for the simple reason that if neither endorsement nor assertion is sufficient to ensure that the postulate of \( T \) is true, then certainly nothing weaker will do. Rather, what’s at issue is precisely whether or not requirement (i) is satisfied given some articulation of \( T \), not whether all conjuncts of \( R \) must be true for the technique to work—that much was already acknowledged. And any such answer or decision that it is satisfied will involve, a fortiori, determining the truth-values of the conjuncts of \( R \). Again, this isn’t to deny that the distinction is a legitimate and important one to make; rather, the point is

\(^{10}\) Thanks to a referee for prompting closer scrutiny of this distinction.
that making it fails to provide functionalists and other enthusiasts of Ramsification with the intended dissolution.

But third, don’t both of these responses fail to abide by basic principles of charity? Surely the theorist marshalling \( T \) will assume that the conjuncts of \( R \) are indeed true and that there is indeed something so denoted; in kind, we should assume that she (literally) knows what she is theorizing about, in so far as she is an expert on the domain thereof. Fine, but goodwill must not cede to bad faith; for these are precisely the assumptions upon which the success of theorists’ implicit definitions turn, and nothing but question-begging is accomplished by mere stipulation or further meta-assumptions about the theorist’s assumptions. We equip ourselves with standards such as truth, not as some sort of theoretical hood ornament, but in order to test whether they’re met; and we do so test precisely to avoid just these kinds of avoidance tactics.

Fourth, one might take issue with the charge of question-begging, arguing that it actually goes the other way around: locutions like ‘determining that all conjuncts of \( R \) are indeed true’ are simply paraphrases of locutions that invoke verification, and it’s this verification requirement that foists the problem onto functionalists—not the technique of Ramsification itself. In response, it’s hardly clear why determination-locutions need be explicated in terms of verification, much less why they need to be explicated at all: in the context of theories of truth, such locutions will just be \( o \)-terms. But even if such explication were necessary, this putative dissolution is only worth taking seriously to the extent that the counter-charge of question-begging is warranted. But what, exactly, has been begged? That Ramsification involves analysing the postulate of \( T \) as a true conjunction \( R \)? That we count implicit definitions of some \( \tau \)-term as successful to the extent we can introduce them in the context of what we know to be true? That verification is a truth-tracking procedure? Surely not the non sequitur that building a verification requirement into Ramsification entails verificationism about truth. Since the reasons for thinking that truth is a standard against which we should test our theoretical principles, it stands to reason that the effect of paraphrasing determination-locutions as verification doesn’t so much undermine the problem of epistemic circularity as amplify it.

Fifth, suppose we instead concede that the credibility of \( T \) does ride on evaluating the conjuncts of \( R \), and that the provision of such evaluation is indeed given by testing for certain standards. Still, one might argue that these requirements are trivially satisfied precisely because the postulate of \( T \) was always postulated to be an ideal theory. There is no better response to such a claim than Russell’s: ‘the method of “postulating” what one wants has many advantages; they are the same as the advantages of theft over honest toil’ [1919: 71]. What does ‘ideal theory’ refer to if not just a theory that is, minimally, true of the phenomena within its scope? And if there is some conception of truth tacit in the conception of an ideal theory, then the extent to which it must be paraphrased away is the extent to which the aforementioned requirement (ii) cannot be satisfied trivially. Hence, we should wonder whether the stipulation is even efficacious to begin with.
The objector could instead follow Lewis, who declared, ‘Suppose the best scientific explanation we can devise [...] includes $T$’ [1970: 428]. Unfortunately, this leaves open the possibility that our ‘best explanation’ is the best of a bad lot, not to mention that ‘best’, like ‘ideal’, would need to be explicated in a way that does not just further conceal some conception of truth.

Alternatively, one might take the concession in an opposing direction, arguing that the requirement (ii) can be read merely as a requirement to determine whether some standard weaker than truth is met. For example, it’s plain that the input over which theorists Ramsify should adequately describe whatever it purports to be about; hence, one might argue that functionalists need only determine that conjuncts of $R$ meet this standard—what Hempel called descriptive adequacy—not truth. Perhaps there’s something to this dissolution, though some clarification would still need to be given as to how descriptive adequacy isn’t just tantamount to truth here (as Hempel thought), as well as why weaker standards like being adequate will suffice when stronger ones like truth and idealization will not.

Seventh (and relatedly) one might again concede that functionalists face a problem of circularity, but argue that the circularity is semantic, not epistemic; for it concerns what functionalists must be able to conceive of or understand in order to know that $T$ is true, not what they must know. Of course, this is not to say that semantic circularity is desirable. As David [2002: 164] rightly put the point, only an empty definition would result from invoking all the facts about alethic, semantic, and pragmatic phenomena, including the facts that involve phenomena related to truth; for any implicit definition of true succeeds when explicitly based on the facts involving truth. However—and here’s the twist—the elimination of semantic circularity from the definiens is already built into the technique of Ramsification.

There’s little reason to think that the epistemic circularity must be redefined semantically, which is reason enough to think that this seventh attempt at dissolution will ultimately be unsuccessful. The objector supposes that functionalists won’t be able to detect whether, e.g., it’s true that true sentences have false negations unless they already understand the meaning of true (or assertion, falsity, negation, etc.). Perhaps, but understanding the meaning of true won’t suffice for ensuring that the implicit definition is successful since it won’t suffice for ensuring that falsehoods aren’t erroneously screened into the conjunction $R$. What is required is that theorists know actual truth-values, not comprehend

---

11Thanks to a referee for raising this clever objection.
12Functionalists need not—and don’t appear to—Ramsify over meaning postulates of the form $x$ means $y$ or explicit definitions (recall, for instance, that they arrive at $T_n^0$ as a result of Ramsification, not prior to it). But they certainly could, and certainly should if moved to make intra-theoretical refinements over time. Suppose that functionalists engage in successive iterations of Ramsification over $T$ (i.e., Ramsification over $T$ on the first iteration, over $T$’s successor $T'$ on the second, over $T'$’s successor $T''$ on the third, etc.). Then we should expect that any claims about truth derived from $T$ on the 1st iteration, such as $T_1^0$, would be acceptable for invoking in successive iterations of Ramsification over any successors of $T$, and that failure to explicitly affirm the claims arrived at when articulating $T''$, $T'''$, etc. would justify doubt about the truth of the initial theory $T$. 

---
possible meanings. Further, given that some principles of $T$ may be non-analytic or *a posteriori*, as functionalists themselves aver, merely understanding what *true* means will not settle whether they’re true: the world itself must cooperate. Further still, whether it’s true that, e.g., truth is correspondence to the facts and not identity to the facts, or *vice versa*, or neither, is not settled by simply accessing a lexical entry for *true* or having some minimal concept *truth*. Hence, with the objector, we can agree that needing to possess the concept *truth*, or understand the meaning of *true*, in order to produce a theoretical definition would be a kind of conceptual, or semantic, circularity distinct from the kind of epistemic circularity involved in having to know that the conjuncts of $R$ are true. Nevertheless, determining that the antecedent of the Carnap sentence with uniqueness holds requires determining the truth-values of the conjuncts therein. So the problem remains.

It seems to be a condition on the possibility of theorists knowing what truth is that they already be in possession of a working concept of truth-values, which presumably includes being in possession of a minimal concept *truth*. An eighth attempt at dissolution takes aim at that claim: surely, it’s possible to possess a minimal concept of $x$ without knowing anything more substantial about what $x$ is a concept of. For example, one can possess the concept *water*, or the concept *sugar*, without being able to determine what water, or sugar, consists in. In particular, one need not know in advance or be able to determine that samples of water are samples of $\text{H}_2\text{O}$, $\text{H}_3\text{O}$, deionized, etc., or that samples of sugar are samples of $\text{C}_{12}\text{H}_{22}\text{O}_{11}$, $\text{C}_6\text{H}_{10}\text{O}_5$, etc., in order to apply *water* to water, or *sugar* to sugar. Rather, someone possesses the concept *water* only if they can, e.g., discriminate amongst samples of water from non-water, recognize this sample of water as being qualitatively similar to that sample of water, etc. Analogously, one can possess the concept *truth* without knowing in advance what realizes the role of truth: ‘it is unsurprising’, wrote Lynch, that we are familiar and adept with the job of truth ‘while remaining clueless about its nature’ [2005: 29]. Therefore, goes the objection, the problem of epistemic circularity rears only because of a faulty premise involving referential opacity.\(^{13}\)

Concepts are simply mental representations of categories: a concept *fence* is a mental representation of fences; *telephone* is a mental representation of hand-held devices used for distal communication; etc. So, being in possession of a concept $x$ involves being able to mentally represent a category of $x$s. But nothing about the problem of epistemic circularity requires denying that knowledge of $x$ can be prised apart from knowledge of the deeper metaphysical features within which membership in the category of $x$s consists. Hence, the objection is based on a misunderstanding about the very problem of epistemic circularity. There are two individually necessary and jointly sufficient claims that must be, and have been, established in order to show why the problem is a problem for theories of truth. One must establish that theorists using Ramsification must

\(^{13}\)Thanks to Michael Horton for pressing this objection.
determine—in whatever epistemic sense anyone wants to help themselves to—that all conjuncts of $R$ are true. Then, one must establish that so determining involves, e.g., already possessing the concept truth, or knowing in advance what truth is. Again, we assume that functionalists and other truth-theorists have some minimal sense of what they are (literally) talking about; were that not the case, the process of amassing the principles constitutive of $T$ could never have begun. Yet, it hardly follows from the claim that theorists need to know in advance what truth is that they need to know in advance what truth consists in (i.e., what its nature is, what its realizers are, whether it is reductively analysable as weak homomorphism for a given sector of discourse, etc.).

Lastly, one might point out that sentences derived by Ramsification may be acceptable for establishing implicit definitions even if the truth-values of some conjuncts of $R$ cannot be determined, or else are determined to be gappy or even just plain false. This point is correct as far as it goes, though it goes against nothing I’ve argued. Lewis [1970: 432] put the point by remarking that unique near-realization might be unique realization near enough: if some $n$-tuple uniquely near-realizes $T'$, then we may be able to correct or slightly modify $T'$ such that the $n$-tuple uniquely realizes its successor $T''$. Bedard [1993: 504–5; see also e.g., Field 1973; Hawthorne 1994] further noted that the $t$-names from a nearly-realized or even unrealized theory partially denote, to varying degrees, each of the $n$-tuples that are fully denoted according to the uniquely realized subtheories of $T$. Yet, we cannot apply these or other suppositions to the case of truth without thereby playing a shell game; for the problem of epistemic circularity’s dogging functionalism will merely arise for the successor $T''$ instead, or for the subtheories of $T'$. Specifically, we might suppose, with Lynch [2004: 405–6], that implicit definition of $t$-terms tolerates a great deal of permutation in the constituents of $T$ so long as there is some essential core or weighted subset—the so-called minimal core of the $F$-role. Yet, this relaxation from all conjuncts of $R$ to some weighted subset $R'$ raises the twin questions of which conjuncts are incontrovertible and why. In answering ‘all the true ones’, we go around in yet another very small circle. And like-minded answers don’t dissolve the problem so much as just repackage it: $R'$ is a smaller sentence with a greater probability of success, but the truth of its conjuncts must still be secured if the definition is to be successful.

6. A Putative Solution

The problem of epistemic circularity calls for positive solutions, not negative dissolutions. I now turn to one, which begins with the following

---

14Pace Field [1973], the inability to determine that all conjuncts of $R$ are true merely leaves the $t$-terminology referentially indeterminate, in which case the definiendum may partially denote more than one thing. Presumably, pluralism about truth is consistent with such results.

15The questions are made even thornier to the extent that error theories, which suggest $T$-schema instances are false, are plausible. (Ironically, some advocates of error theories have employed functionalism about truth e.g., Devlin [2003], to support their view.)
reminder. Knowledge of what the τ-term true refers to is not common knowledge, *ex hypothesi*, given what it means for a term to be a τ-term. Consequently, it would be a mistake to treat the functionalist’s thesis,

\[(\text{fact}) \ \text{true} \text{ refers to the single higher-order functional role property } F \text{ of having lower-order properties } \rho_1, \ldots, \rho_n \text{ that realize it,}\]

as a description of common knowledge—a claim that’s also supported by the sole empirical study on the matter [Næss 1938]. And so, it stands to reason that the expression of what is common knowledge needn’t—indeed, shouldn’t—be the very same τ-term under analysis.

As a solution to the problem of epistemic circularity, this proposal is satisfied easily. The truth-theorist need only invoke some o-term, say *true, that expresses the ordinary street-level concept, *truth, where the latter is the mental representation of some category structured by the property *true. In doing so, she can show that her employment of Ramsification merely involves an implicit commitment to the *truth of all conjuncts of \(R\). Any further evaluation of the conjuncts of \(R\) merely involves explicit decisions that they are *true, relative to how the world is. And such decisions rest simply on one’s basic linguistic competence with the o-terms and o-sentences of \(L\). Problem solved.

This positive solution has several happy consequences. First, it accords well with many functionalists’ specific emphasis on common knowledge and our folk understanding [Lynch 2000, 2004, 2005]. It also accords well with the general motivation to use Ramsification; for the point was always to use old, ordinary, and original terms against which the τ-terms inaugurated by a theory can be holistically situated. And it also vindicates functionalists’ use of Ramsification as, again, something more than a gratuitous logical drill. Under this solution, Ramsification proceeds in the usual way. Interdefinable τ-terms are replaced with existentially bound variables; and since *true is an o-term, no variable need replace it. Presumably then—or, at least, under the canonical Lewisian version—semantic circularity is avoided. And no epistemic circularity ensues since alethic theorists implicitly define true by explicitly helping themselves to the property referred to, and concept expressed, by *true. Best of all, Ramsification itself is not needed to define o-terms like *true that issue from a commonsense conception or street-level vocabulary, since Ramsification is not needed to define o-terms.

One consequence, however, may seem less happy. This solution requires of functionalists that they give up or otherwise tread on their monist assumptions. Why? Because appealing to the property *true to evaluate each conjunct of \(R\), such that Ramsification can be used to implicitly define a technical or specialized τ-term true that refers to a single domain-general higher-order multiply-realizable functional role property \(F\), involves proliferating properties in exactly the way that functionalists meant to avoid. Call this the pluralist’s revenge.
7. Remarks on the Pluralist’s Revenge

7.1 Initial Objections Addressed

Some questions and objections can be immediately dispensed with. For instance, isn’t pluralism about truth defective, given that, pace Lewis, unique realization is the standard of correctness to insist on? Not necessarily. If \( T \) is a version of pluralism about truth, then presumably we’d count it as being true only if more than one ordered \( n \)-tuple of entities satisfies its realization formula, and false otherwise. So unique realization is precisely the wrong standard to insist on for such a theory, and we ought not to treat pluralism, if true, as defective simply for describing what, in that case, is the case. Either way, whether pluralists have a viable position themselves is an open question that’s in little need of being settled here.

Other questions arise for which no simple or definitive answer is available. For example, what features do \(*\text{truth} \) and \( *\text{true} \) have, and are those features really sufficient to do the work asked of them? Here, the best one can do is to note several constraints that would need to be satisfied for the solution to be viable, and to argue that \(*\text{truth} \) and \( *\text{true} \) have the features that do satisfy these constraints. For instance, one constraint is that \( *\text{true} \) and \( \text{true} \) must be referential without being co-referential; another is that their referents should be a species of accuracy or correctness (rather than legitimacy, authenticity, fidelity, etc.). Prima facie, we have no reason to suppose that such constraints aren’t satisfied, and that’s reason enough for supposing that they are. But ultimately, whether \( *\text{true} \) is a species of accuracy or correctness but not co-referential with \( \text{true} \) depends, in part, on how speakers of \( L \) use the \( o \)-term \( *\text{true} \). And that’s simply an empirical matter for the lexicographer.

Other questions and objections, however, cannot be addressed obliquely. For example, suppose functionalists complain that the predicate \( \text{is } *\text{true} \) was introduced stipulatively as a ‘designer’ predicate—i.e., one introduced by design solely to solve the problem of epistemic circularity. It’s unclear why anyone should accept that \( *\text{true} \) was introduced stipulatively—ex hypothesi, it was not. But, even allowing that it was, the complaint nevertheless confuses the introduction of \( *\text{true} \) stipulatively as an \( o \)-term, which entails only that it has achieved unit status in the grammar of \( L \) prior to the formulation of \( T \), with the introduction of it stipulatively into the grammar of \( L \). So, if \( *\text{true} \) was introduced stipulatively, then it was introduced stipulatively as an \( o \)-term, and so the complaint couldn’t be that it isn’t, though it could be revised to be that \( *\text{true} \) cannot possibly be an \( o \)-term, and so cannot be stipulatively introduced as one.

The revised complaint would be reasonable if there were evidence for thinking either, first, that \( *\text{true} \) has achieved unit status in the grammar of \( L \) but not prior to the formulation of \( T \), or else, secondly that \( *\text{true} \) just hasn’t achieved unit status simpliciter. In neither case do we need to pass the issue off as an empirical question for the lexicographer. The first thought
presupposes that *true is indeed a term expressing the corresponding street-level alethic concept possessed by ordinary folk, and it’d be absurd to suppose that such folk only expressed themselves thus following the formulation of functionalism about truth. But if we take *true to express the corresponding street-level alethic concept possessed by ordinary folk, then the only warrant for the second thought that it hasn’t achieved unitization would be that *true is nowhere apparent in spoken or written discourse. And this belies a fundamental confusion about the notation *true. Neither the solution nor the pluralist’s revenge entails the claim that ordinary folk mark the expression of their street-level alethic concept differently than truth-theorists’ use of true. Quite the contrary; it’s precisely because the phonological or orthographic representation of the o- and t-terms do converge—e.g., [troo] and [true], respectively, in English—despite the fact that the semantic representations diverge, that there’s need, here, for some diacritic or other convention to distinguish the o-term that ordinary folk do use from the functionalists’ technical term true.

7.2 A Deeper Challenge

Both the solution and the pluralist’s revenge raise an important series of questions about the relationships (e.g., hyponymy, polysemy, homonymy) between technical terms introduced by theories of truth and their lay counterparts. The proposal thus far has been that, for any given theory of truth T, terms like truth and true are its theoretical terms (what else could they be?), and that, as an o-term, *true is pre-theoretical in the straightforwardly literal sense that it has unit status in the grammar of L prior to the formulation of T. As an o-term, *true may also occur in (T: P1, . . . , Pn), though it need not—nothing about Ramsification requires that all o-terms be included in (T: P1, . . . , Pn), just as nothing about Ramsification requires that the articulation of any given theory includes a dictionary. On the other hand, truth must occur in (T: P1, . . . , Pn), since it’s the target t-name replaced by a variable in deriving T' or any successor formula.

One interesting manoeuvre might be for functionalists to argue that it’s a mistake to think *true is always and everywhere an o-term. Unitization in the grammar of L prior to the formulation of T doesn’t rule out that *true is an o-term relative to T but a t-term relative to the formulation of some earlier theory S. According to Lynch [2004, 2005], speakers of L do possess just such a universal folk theory of truth S, whose principles (O1, . . . , On) express the ordinary alethic concepts that give rise to L and its metalinguistic extension. In particular, *true is a t-term issuing from this universal folk theory, in which case—goes the objection—functionalists’ implicit definition of true merely captures that folk-theoretical use metatheoretically.

This manoeuvre aims to help functionalists maintain their monistic assumptions by showing that there’s exactly one truth property. Its success depends on functionalists’ being able to show that the principles fixing their
τ-term true and folk theorists’ τ-term *true turn out to be the same or substantially overlapping—i.e., the conjecture that \( T \cong S \). After all, only in that case could the two properties referred to by true and *true be reductively identified, and only with such an identification would it have been shown both that there’s exactly one truth property and thus that the pluralist’s revenge is avoided.

A litany of problems accompanies any such reductive identification, however. At best, this manoeuvre is only partially successful; for even if *true is indeed an o-term relative to \( (T: P_1, \ldots, P_n) \) but a τ-term relative to \( (S: O_1, \ldots, O_n) \), both the solution and the pluralist’s revenge remain unaddressed in the case where *true is an o-term relative to \( T \). And where *true is a τ-term relative to \( S \), the manoeuvre may not even be partially successful; for if implicit definition by Ramsification over \( (S: O_1, \ldots, O_n) \) is needed, the problem of epistemic circularity would simply reoccur. The potential for regress ought to be avoided where possible.

The problems worsen. Even if a reductive identification were fully successful, the problem of epistemic circularity would then just be reinstated; for the solution depends on there being some numerically distinct property *truth had by all conjuncts of \( R \), such that it’s sufficient for defining the τ-term true that the theorist determine the *true sentences in which it features as a subsentential component. Moreover, any such reductive identification would also seem to be self-abdicating; for it would entail that the functionalist’s property truth just is the good old-fashioned property referred to by ordinary folk. But in that case, it’s unclear why anyone would ever need a functionalist theory of truth employing Ramsification to implicitly define it as such in the first place.\(^{16}\)

Clearing these obstacles brings us to the deeper issue of whether the conjecture \( T \cong S \) is gripping, and, if so, then to what extent their principles overlap. There’s reason to suspect substantial differences. First, it’s quite clear that functionalism about truth is a technical philosophical theory \( T \), along the lines of \( (\text{infl}) \), about a specific functional kind \( F \) to which truth purportedly reduces. In the context of such a theory, the τ-name truth is introduced to denote it. In the context of a substantially different philosophical theory—e.g., a correspondence theory postulating a two-place structural relation to ‘Cambridge’ facts—it’s also clear that implicit definition will yield a different meaning and denotation, since, obviously, the theoretical principles of functionalist and structuralist theories differ substantially. And the substantially different theoretical principles of either of these technical theories will differ substantially from those of the universal folk theory \( S \). Functionalists themselves have made this latter claim a central part of their view:

\[ [\text{The principles we employ in our folk theory are those the folk tacitly believe, or are rationally committed to. They aren’t those principles that result from technical philosophical argument: thus principles that concern the nature of} \]

\(^{16}\)Also unclear would be how to uphold the very distinction between what Carnap and Lewis called the \textit{analytic versus synthetic postulate} of \( T \) (i.e., τ- and o-sentences, respectively) if o-terms and τ-terms are sometimes co-referential, synonymous, or otherwise interdefinable.
correspondence, reference, coherence, superassertibility, and the like are not part of our folk theory.

[Lynch 2004: 393 n. 6]

Conversely, the folk-theoretic platitudes or truisms of $S$ typically don’t appear in technical philosophical theories $T_1, \ldots, T_n$, and typically don’t do any heavy lifting if and when they do, as Chomsky observed.17

A good part of contemporary philosophy of language is concerned with analyzing alleged relations between expressions and things, often exploring intuitions about the technical [terms] denote, refer, true of, etc. said to hold between expressions and something else. But there can be no intuitions about these [terms], just as there can be none about angular velocity or protein. These are technical terms of philosophical discourse with a stipulated sense that has no counterpart in ordinary language.

[2000: 130; cf. Næss 1938]

But now notice that if Lynch and Chomsky are correct about the distinction between technical philosophical theories and non-technical folk theories, then making good on the claim that true and *true are co-referential would require ignoring non-trivial and possibly vast differences between $(T: P_1, \ldots, P_n)$ and $(S: O_1, \ldots, O_n)$.

In the absence of a bijective function mapping functionalists’ principles $(T: P_1, \ldots, P_n)$ into $S$, then we have numerically distinct conjunctions, and thus numerically distinct truth-conditions, associated with each postulate; and the upshot is the positing of numerically distinct, even if only partially denoted, $n$-tuples. After all, recall that the cardinality of the extension of truth is determined by the number of unique functional specifications, where each unique functional specification implicitly defines a distinct $\tau$-term. Functionalists about truth must therefore take care to not proliferate functional specifications. Yet, functional specifications are easy to proliferate—formally speaking, one need only begin altering the identity- and truth-conditions of the content constituting them [Wright 2005]. Of course, as Lynch rightly pointed out, ‘slight differences in [content] do not entail that we are simply talking past one another’ [Lynch 2005: 39]. We needn’t disagree with that point, however, since the relevant differences under consideration between alethic theorists and ordinary folk were never the slight ones to begin with; rather, they are increasingly tending towards a yawning conceptual gap. In any case, suffice it to say that there are sobering difficulties with functionalists’ showing that pluralists are wrong to proliferate truth properties, as there

17Lynch, Horwich, and others often take the principles of a theory of truth $(T: P_1, \ldots, P_n)$ to be platitudes, where platitude picks out those claims that, inter alia, reflect the ordinary uses of terms in a given linguistic community. Yet, $\tau$-terms have no such ordinary uses for platitudes to reflect, ex hypothesi. Therefore, $\tau$-terms do not occur in platitudes. Platitude-talk therefore seems to instigate a dilemma. If platitudes are the constituents of $T$ and true is a $\tau$-term, then true cannot occur in $T$ and so cannot be implicitly defined by Ramsification over it. But then, the only terms whose ordinary uses could be reflected by platitudes are $o$-terms; and so if true is instead an $o$-term, then it can occur in $T$. But then Ramsification would again become superfluous, since Ramsification is unnecessary for defining $o$-terms.
are non-trivial difficulties in identifying and individuating the $F$-role monistically.\textsuperscript{18}

California State University, Long Beach

Received: January 2008

Revised: December 2008

References


Pedersen, Nikolaj 2006. What can the Problem of Mixed Inference Teach us about Alethic Pluralism?, \textit{The Monist} 89/1: 103–19.


\textsuperscript{18}I would like to thank Thomas Sattig, Michael Horton, and two anonymous referees for their constructive comments.