Scott Soames: *Understanding Truth*

MATTHEW MCGRATH
*Texas A & M University*

Scott Soames has written a valuable book. It is unmatched in its clear expositions and evaluations of the theories of truth of Tarski, Kripke, and Strawson, and much other formal and philosophical work on truth. It also incorporates ideas from several of Soames’ published papers on the subject. Its centerpiece, however, is the application of a theory of partial definitions to the liar paradox and the problem of vagueness. Most of my discussion will focus on this important contribution. I close with an examination of Soames’ discussion of T-sentences.

I. The Liar Paradox and Partial Definition

Consider Soames’ Version 1 of the liar paradox.¹

i. ‘(1) is not true’ is true iff (1) is not true

ii. (1) = ‘(1) is not true’

iii. (1) is true iff (1) is not true

iv. (1) is true and not true.

We stipulate that (ii) is true. Soames’ response is to reject a presupposition (EM) of the argument

(EM) (1) is either true or not true

Which step merits rejection, then, depends on whether we read ‘iff’ throughout as the material biconditional. If we do, then (i) reduces to a disjunction

Either ‘(1) is not true’ is true and (1) is not true or ‘(1) is not true’ is not true and (1) is true.

¹ p. 50. All references are to Soames’ *Understanding Truth* (New York: Oxford University Press, 1999).
Since we should reject both (1) and its negation, we should reject both disjuncts in the above disjunction. Thus, we should reject (i). If we don’t read ‘iff’ materially, then we will read A iff B as true iff A and B have the same status, true, false, or undefined. In this case, we ought to accept (i)-(iii) but reject (iv), since to reach (iv) from (iii), we need the problematic (EM). So whichever way we read ‘iff’, we block the conclusion.

How can we reasonably reject both a sentence and its negation? Soames suggests that the standard reasoning used in explaining the paradox justifies the rejection of (1) and its negation. This reasoning consists in first supposing (1) is true and concluding it’s not, and then supposing that (1) is not true and concluding that it is. In a later chapter, Soames explains more fully how both rejections are justified by appealing to the theory that ‘true’ is a partially defined predicate. I will discuss this theory later, but first I want to ask about the implications of rejecting (1) and its negation.

One might hope to avoid rejecting these sentences by dismissing them instead. Dismissal is the attitude one would take when, out of the blue, your companion cups thin air and with seeming seriousness makes the remark, “This is a fine red one.” One would not reject the remark, for it fails to express a proposition in the context. However, Soames allows that (1) and its negation, as well as (EM) and its negation, all express propositions. Dismissal appears out of the question.

If I reject ‘p’, what have I done? I have done more than merely not accept ‘p’. Presumably, if I am sincere, I have expressed my disbelief that p. But what grounds are there for distinguishing a disbelief that p from a belief that not-p? If there were no distinction, then if I were to reject both ‘p’ and ‘not-p’ sincerely, I would be expressing beliefs in contradictory propositions.

Yet, one also wonders whether, even if disbelief is distinct from belief in the negation, disbelieving a proposition and its negation isn’t just as bad as believing them, especially if we know what we’re doing. Doesn’t disbelief, like belief, have a “mind-to-world” direction of fit? Isn’t a disbelief that p correct if not-p, just as a belief that p is correct if p? If so, then in disbelieving a proposition and its negation, I guarantee that one of the two disbeliefs is incorrect. Presumably, guaranteeing one’s own incorrectness (regarding such a small number of propositions) is part of what is so bad about believing at once both a proposition and its negation.

To save words, I call the unnegated ‘p’ the negation of ‘not-p’.

See p. 166, p. 177, p. 177n16. Soames uses ‘≡’ to express this biconditional.

p. 159.

This is a slight revision of Strawson’s example, used by Soames. See p. 168.

In his reply, Soames points out a slip in this reasoning. Without excluded middle, the conclusion I draw here requires the conditional: if it’s correct to disbelieve that p, then not-p. Let me correct this slip. The needed conditional is derivable from the innocuous claim, referred to in the text, that if p, then it’s correct to believe that p, together with principle (ii) of Soames’ footnote 4, which Soames does not challenge. Thus:
Might we, rather than rejecting or dismissing the relevant sentences, simply withhold on them, i.e., refuse to accept, reject or dismiss them? One problem, mentioned earlier, with this option is that it conflicts with our sense of what the liar-style reasoning shows about (1) and its negation. Generally, if the supposition of ‘p’ leads to ‘not-p’, then it leads to a contradiction, ‘p & not-p’, which must be rejected, and so we must reject also the supposition ‘p’.

So, one worry I have about Soames’ response to Version 1 of the liar paradox concerns the propositional attitude associated with rejecting a sentence. A second worry concerns logical inference. Recall that we are reading the ‘iff’ in (iii) as Soames’ ‘='. Consider the corresponding conditional, which I express by ‘→’. Soames clearly must reject the following sequent:

\[ p \to q \vdash p \supset q \]

But how can he? Every conditional obeys *modus ponens*. Now assume ‘p→q’ as a premise and assume ‘p’ for conditional proof. We reach ‘q’ through *modus ponens* for ‘→’, and discharging ‘p’, arrive at ‘p \supset q’. Alternatively, we may use RAA on ‘¬(p \supset q)’, use a derived rule linking ‘⊨’ with ‘&’, and then apply *modus ponens* for ‘→’ to reach ‘p \supset q’.

Conditional proof and RAA turn out to be invalid, given Soames’ semantics, for they sometimes lead from true premises to an undefined conclusion. What this means is that Soames must give up important parts of the proof theory of classical logic (even if he does preserve the classical semantics of ‘and’, ‘or’, ‘not’, ‘some’ and ‘all’). I point this out only because, to my knowledge, it is not made explicit in the book.

I now turn to the theory of partially defined predicates. Soames explains the basic idea by introducing one himself. Suppose we have two groups of people, A and B. Those in group A are quite short and those in group B are of moderate height, and no one in group A is as tall as anyone in group B. Let us then say:

\[ a. \text{If it's correct to disbelieve that } p, \text{ then it's not correct to believe that } p \text{ (this is equivalent to Soames' (ii))} \]
\[ b. \text{If it's not correct to believe that } p, \text{ then not-}p. \text{ (this is equivalent to the innocuous claim)} \]

Therefore, c. If it’s correct to disbelieve that p, then not-p.

---

7 See again p. 159.
8 On this point, see p. 208.
9 Otherwise, he will have to accept (iv), given his acceptance of (iii).
10 At least, every conditional obeys *modus ponens* restricted to cases in which neither the antecedent nor the consequent is or contains a conditional.
11 This discussion of the properties of ‘→’ may be beside the point, since (EM) itself is a theorem of classical logic, provable using RAA, “or” introduction, and the commutativity of “or”.

412 MATTHEW MCGRATH
i. Every member of group A is a smidget. Further, for any adult, if his or her height is less than or equal to the height of at least one member of group A, then s/he is a smidget.

ii. Every member of group B is not a smidget. Further, for any adult, if his or her height is greater than or equal to the height of at least one member of group B, then s/he is not a smidget.

iii. Nonadults (and nonhumans) are not smidgets.12

This partial definition gives sufficient conditions for being a smidget and for not being a smidget, but it doesn’t give us necessary and sufficient conditions. This means that for certain objects the predicate ‘smidget’ is undefined. And if we know that ‘smidget’ is undefined for an object, we should reject both \[ \neg \alpha \text{ is a smidget} \] and its negation, where we use \( \alpha \) to name the object.

One might think that if ‘smidget’ is undefined for an object, then it does not apply to it, and that, knowing this, we must accept the sentence that says it doesn’t. But this follows only if we take ‘applies to’ to be fully defined. Soames maintains instead that ‘applies to’ is partially defined as well:

The predicate ‘red’ applies (does not apply) to an object \( \equiv \) it is (is not) red.

The predicate ‘smidget’ applies (does not apply) to an object \( \equiv \) it is (is not) a smidget.

(and so on, one clause for each predicate in the language.)13

If we reject both \[ \neg \alpha \text{ is a smidget} \] and its negation, then we must reject both \( \neg \text{‘smidget’ applies to } \alpha \) and its negation. If we partially define ‘true’ for atomic sentences in terms of predicate application, then we will reject both \[ \neg \text{‘is a smidget’ is true} \] and its negation.

Soames employs an account along these lines, invoking the apparatus of Kripke’s theory of truth, to justify the rejection of the instance of Excluded Middle (EM) above. Our linguistic conventions regarding ‘true’ leave (1) (= ‘(1) is not true’) and its negation, and so also (EM) and its negation, undefined. So these sentences must be rejected.

I have two concerns about this. First, it is unclear to me why we should reject \[ \neg \alpha \text{ is a smidget} \] and its negation, given Soames’ definition. Soames might reply that we can’t dismiss them as we can in the case of ‘This is a fine red one’, because they express propositions. But focusing on an object o

---

12 p. 164.
13 p. 166.
named by $\alpha$ for which we know ‘smidget’ is undefined, it seems more likely that we would rather withhold judgment on $\neg \alpha$ is a smidget $\neg$ and its negation. Our definition, after all, is partial; it says nothing about such cases. We would be going beyond that definition to reject the claims in question. Imagine we later decided to include some previously undefined cases in the extension of ‘smidget’. Would we have to regard ourselves as then making claims we previously rejected, or rather as making claims on matters about which we previously were silent? However, Soames needs our attitude in the case of (1) to be rejection. For withholding here wouldn’t justify rejecting the contradiction “(1) is true and (1) is not true”.

This reveals a key difference between ‘true’ and ‘smidget’. As Soames points out, there is no way of augmenting the linguistic conventions for ‘true’ to consistently define ‘true’ for (1). But there is a way of augmenting the conventions for ‘smidget’ to define any previously undefined case. This may explain the fact that we (or at least I) incline toward rejecting (1) but toward withholding on $\neg \alpha$ is a smidget $\neg$.

My second concern is about the proposition $\neg \alpha$ is a smidget $\neg$ expresses and how it gets to express it. We may say it expresses the proposition that $\alpha$ is a smidget, but what proposition is that? Consider the thinkable thoughts prior to the introduction of ‘smidget’. With its introduction, does Soames create a new range of thinkable contents, propositions? That seems implausible. If ‘smidget’ latches on to a property, it latches on to one that existed prior to the introduction of the word. Soames denies it latches onto a “totally defined” property, that is, a property that either holds or fails to hold of every object. So he concludes, in effect, that it must latch onto a “partially defined” property. But how does it do that? The partial definition of ‘smidget’ does not fix uniquely on a totally defined property, since there are many such properties that satisfy the conditions given. Yet it isn’t clear how it could fix uniquely on a partially defined property, either. If ‘smidget’ fixes on a property, it does so in virtue of the property’s satisfying the conditions stipulated above. But these conditions don’t fix on any unique property, whether totally or partially defined.

One might, in any case, worry about the philosophical consequences of appealing to partially defined properties. Properties, unlike predicates, do not exist and have application conditions by courtesy of our linguistic conventions. But, if Soames is right, some of them are partially defined, insofar as they determine only partial application functions. So, partial definition, the phenomenon in terms of which Soames aims to blunt the liar and Sorites paradoxes, is at bottom not a linguistic phenomenon; there is “undefinedness”.

Similarly, one is less tempted to reject (2) (= ‘(2) is true’) or ‘Every sentence written in this paper is true’, than merely to withhold judgment.

What he says is that $\neg \alpha$ is a smidget $\neg$ expresses a proposition that determines only a partial function from circumstances to truth-values. See p. 170.
in the world. This may disappoint those who wish to construe these as problems for the philosophy of language, and not for metaphysics.

In light of such problems, it might be preferable to mimic a supervaluationalist approach, at least to the extent of invoking the notion of partially expressing a property. One would say that ‘smidget’ doesn’t fully express any property, but partially expresses many (totally defined) properties. A property P is partially expressed by a predicate \( \Phi \) iff every object that the linguistic conventions stipulate to be in the extension (antiextension) of \( \Phi \) has (does not have) P. We then define the notion of the determinate extension (antiextension) in terms of partial expression: an object is in the determinate extension (antiextension) of \( \Phi \) iff it has (lacks) every property that \( \Phi \) partially or totally expresses. This strategy frees Soames’ account from its associations with ontological vagueness. And so long as one does not identify truth with determinate truth, one can retain the truth-schema.\(^{16}\)

II. T-Sentences

In a subtle discussion, Soames considers and rejects an argument that homophonous instances of the English T schema are definitional of ‘true in English’. The argument assumes that knowing what “‘---’ is true in English iff --” means in English is enough to know that instances of the English T schema are true. The argument then claims that if, in addition to this, one understands an instance of the English T-schema, then one has enough information to accept it, know it is true, and know the truth of the proposition it expresses. From these considerations, the argument draws its conclusion.

Soames replies that one might understand an instance of the T-schema but not know that what ‘English’ refers to in one’s language is one’s language, even given that English is one’s language. If this is the case, then although I may know what “‘---’ is true in English iff --” means in English and also know that every instance of it is true in English, that isn’t enough information to accept any particular instance e, nor to know of any such e that is true in English, nor to know the truth of the proposition expressed. He offers two possible reasons. First, I may not know that the relevant substitutend in the instance is a normal sentence of the language designated by ‘English’; and second, even if I do know this, I may still not know that the sentence means in that language what it means in my language.

\(^{16}\) What of sentences such as ‘Mary believes that Mr. Smallman is a smidget’, where Mr. Smallman is an undefined case? It would be too much to require Mary to believe each of the propositions partially expressed by ‘Mr. Smallman is a smidget’. As a first approximation, we could say this: ‘A believes that S’ is determinately true iff, there is some proposition that S partially or totally expresses, and for every such proposition P, the referent of A has a belief that has P as partial content. This admittedly involves locating “undefinedness” in thought, with the consequence that “undefinedness” is not, in all cases, a matter of linguistic convention. Nevertheless, the source of undefinedness is representation, not reality.
Three conditions on the definitionality of instances of the T-schema are discussed. Let S be a person who knows what "'---' is true in English iff --" means in English and understands an instance e. The conditions on S and e are these: (a) that S accepts e, (b) that S knows e is true in English, (c) that S knows the truth of the proposition e expresses in English. I want to argue that these conditions can be met if we introduce an intensional language-relativity into the concepts of understanding and accepting a sentence.

Consider the English 'A fortnight is a period of two weeks'. And consider also someone, Bob, who speaks a language Schmenglish, just like English, but in which the meanings of 'fortnight' and 'month' are interchanged. Bob understands 'A fortnight is a period of two weeks', and may even know what it means in English, but he won't accept it. Still, surely, the sentence is definitional of 'fortnight' in English. The problem is that we're not distinguishing between understanding or accepting this sentence as English and as Schmenglish. To understand a sentence as English is to know what it means in English. To accept a sentence as English is to accept it while regarding oneself as speaking English. As English but not as Schmenglish, Bob accepts 'A fortnight is a period of two weeks'. Bob, therefore, is no barrier to the claim that understanding this sentence as English is enough to accept it as English.

Now suppose we reformulate conditions (a)-(c) in terms of understanding and accepting as English. The conditions are then satisfied. If I understand an instance of the T-schema as English, I will accept it as English, know it is true in English, and know the truth of the proposition it expresses in English.17

Given these facts, one is tempted to conclude that the instances of T-schema are definitional of 'true in English'. For such instances are "analytic in English": understanding them as English is sufficient for accepting them as English, for knowing them to be true in English, and for knowing the truth of the propositions they express in English. Even more can be said: understanding "'---' is true in English iff --" as English is sufficient for knowing of an instance that it is true in English, so long as one knows that the substitutend in it is a normal sentence of English. How, then, can such instances be non-definitiona[17]

---

17 The intensionality of talk of "understanding (accepting) as" shows up in the case in which I know
   (i) what "'---' is true in English iff --" means in English
   (ii) that 'Snow is white' is a normal sentence of English,

but misunderstand 'Snow is white' as English, because, say, I believe that it expresses in English the proposition that snow is red. In this case, I understand 'Snow is white' as (a sentence of) my language but not as English, and I accept "'Snow is white' is true in English iff snow is white" as English but not as my language.
Yet they do seem to be so.\textsuperscript{18} A theory of truth invoking propositions, such as the one Soames accepts, can help us here. Suppose I know what "---' is true in English iff ---" means in English and also know that 'Snow is white' is a normal sentence of English. (I may or may not know what it means in English.) Then I know that the right side of "'Snow is white' is true in English iff snow is white" is true in English iff the proposition P that 'Snow is white' expresses in English is true. I also know that the left side is true in English iff what 'Snow is white' expresses in English is true (since that is what the left side says in English). But what 'Snow is white' expresses in English, we've seen, is P. So both sides are true in English iff P is true. Thus, the left side is true in English iff the right side is. I therefore know, merely on the basis of such knowledge of English, that the biconditional is true in English. And I establish this using a theory of sentential truth for which the instances of the T-schema are not definitional of the truth-predicate, viz. the theory that a normal sentence of a language L is true in L iff what it expresses in L is a true proposition. If I then come to understand 'Snow is white' as English, I will accept its T-biconditional as English and know the truth of the proposition it expresses in English.

\textsuperscript{18} As Soames points out, a person could grasp the proposition such an instance expresses and doubt it, as in the case of a monolingual French speaker asking herself (in French) whether 'Snow is white' is true in English iff snow is white.