

Metarepresentation and the Cognitive Value of the Concept of Truth^{*}

Gurpreet Rattan
University of Toronto

This paper explicates the concept of truth in terms of the notion of cognitive value at three interconnected levels of increasing concreteness. The first level is programmatic and exploratory, aiming to illuminate both the concept of truth and the notion of cognitive value. The second level of explication is dialectical, aiming to reveal a blind spot in a broad range of contemporary views of the concept of truth. The third level of explication is substantive, aiming to explain the cognitive value of the concept of truth in terms of its partially constitutive role in critical reflective thinking.

This paper explicates the concept of truth in terms of the notion of cognitive value. The explication proceeds at three interconnected levels of increasing concreteness. The first level is programmatic and exploratory. The aim at the first level of explication is to illuminate simultaneously the understanding of both the concept of truth and the notion of cognitive value by making use of the notion of cognitive value as the fundamental analytical tool in the explication of the concept of truth.¹ The second level of explication is dialectical. At this level the aim is to reveal a blind spot in a broad range of contemporary views of the concept of truth. I will be arguing that neither *deflationary* nor *inflationary* views of the concept of truth assign cognitive value to *directly specified truth attributions* as they are *intuitively understood*. The third level of explication is substantive. At the third level, the aim is to explain what the cognitive value of directly specified truth attributions intuitively understood is. I will be arguing that directly specified truth attributions intuitively understood are partially constitutive of the epistemic advance made in *critical reflective thinking*.

I

How should one inquire about the cognitive value of the concept of truth? I begin with a few rough words about the notion of cognitive value, asking the reader to bear in mind that it is the burden of the paper at the programmatic and exploratory level to say more.

The cognitive value of a concept is the contribution that grasping or understanding the concept makes to a thinker's cognitive or epistemic perspective. In a helpful slogan, if X is a concept of x, then the cognitive value of X is to make possible thinking and knowledge of x as x. Thinking of x as X is to employ a particular *way of thinking* of x. The notions of cognitive value and ways of thinking derive from Frege, with their most famous application occurring at the beginning of 'On Sense and Reference' (Frege 1892) to what is expressed by identity sentences. Cognitive value and

^{*} Thanks to Paul Pietroski and Michael Glanzberg for illuminating discussion, and to Cory D. Wright for very helpful feedback on an earlier draft of this paper.

¹ I use 'concept' and 'notion' not to distinguish kinds of intentional entities (call them all 'concepts'), nor even to distinguish kinds of concepts, but to distinguish what is being used here as a theoretical concept, COGNITIVE VALUE, in a theory that theorizes concepts themselves, from what is here the target concept being theorized about, namely TRUTH. I use 'notion' for the former, 'concept' for the latter.

ways of thinking are what Frege's 'intuitive criterion of difference' or 'differential dubitability test' are intended to track. I will be thinking of the criterion or test as follows:

Criterion of Difference for Thoughts:

The thought p is distinct from the thought q iff someone can without irrationality take conflicting cognitive attitudes to p and q .²

This criterion, like the slogan mentioned above, supplies guidance, and I will be applying it in explicating a broad range of views of the concept of truth. But I will be suggesting that explicating the concept of truth in terms of the notion of cognitive value illuminates the notion of cognitive value by involving applications of the notion that neither conform to the slogan nor are governed by the intuitive criterion.

A first place to turn in explicating the concept of truth in terms of cognitive value is to the use of the concept of truth in cognition and of 'true' and cognates in language. We can begin by familiarly distinguishing between two central kinds of logico-linguistic environments in which one finds the concept of truth (cf. Quine 1970: 10–13; Field 1994: §5; Horwich 1998: 2ff.; Soames 1999: 22–23). The first involves attribution of truth to a representation that is specified in the truth attribution in a direct and explicit way.³ For example

1. 'Snow is white' is true
2. It is true that snow is white
3. The thought that snow is white is true

are examples in which truth is attributed to a representation that is itself represented in highly direct ways—by quotational name, by that-clause, as the thought that p . The second kind of logico-linguistic environment involves attributions of truth to bearers that are mentioned indirectly,⁴ as in

4. What John uttered is true
5. Everything Plato believed is true
6. A conjunction is true iff its conjuncts are true

Here truth is attributed to representations, but these representations are mentioned indirectly by generalization or description. I will call cases of the first sort

² 'The thought associated with one sentence S as its sense must be different from the thought associated with another sentence S^* as its sense, if it is possible for someone to understand both sentences at a given time while coherently taking different attitudes towards them...' (Evans 1982: 18–19). Thought-theoretic analogues of the criterion can be found in Goldberg (2008: 165), Kimbrough (1998: 470), and Peacocke (1992: 2), where Peacocke's 'Distinctness of Concepts' criterion strengthens the criterion into a biconditional. I will be making use of a biconditional version.

³ '[T]he truth predicate is attached to an explicitly articulated proposition' (Horwich 1998: 39).

⁴ '[W]e predicate truth of some proposition or set of propositions that we do not explicitly assert, display, or produce' (Soames 1999: 23); 'what is said to be true is not directly presented or displayed' (Soames 1999: 235).

metarepresentationally directly specified truth attributions and cases of the second *metarepresentationally indirectly specified truth attributions* (DSTAs and ISTAs for short).⁵

If we ask now about the cognitive value of the concept of truth, we are asking about the cognitive value of DSTAs and ISTAs, and about what these cognitive values together determine as the cognitive value of the concept of truth. That is, we are asking:

- Q1: What is the cognitive value of DSTAs?
- Q2: What is the cognitive value of ISTAs?
- Q3: What do the cognitive values of the two kinds of truth attributions determine as the cognitive value of the concept of truth?

My primary (although not sole) concern in this paper will be with Q1 and the cognitive value of DSTAs.

The cognitive value of a concept (or thought) is the contribution to a thinker's epistemic perspective that understanding the concept or thought makes. My primary concern will be the cognitive value of DSTAs. The next section takes up the question of in what understanding DSTAs consists.

II

What we intuitively understand when we understand DSTAs is different, but not entirely different, from what we intuitively understand when we understand ISTAs. Let me use 'representation' to cover bearers of truth in general (sentences, utterances, beliefs, thoughts, propositions, etc.). Our intuitive understanding of both DSTAs and ISTAs involves understanding that truth is being attributed to representations. In both DSTAs and ISTAs the concept of truth combines with a device of metarepresentation to attribute truth *to* representations. But the way representations are represented in the different truth attributions is different.

ISTAs are such that one can understand them without knowing to which representation truth is being attributed. One can understand that what John uttered is true without knowing which sentence John uttered. One can understand that everything Plato believed is true without knowing which thoughts Plato believed. Further, not only can one understand that everything that Plato believed is true without knowing which thoughts Plato believed, one can do so without thinking or entertaining the thoughts Plato believed (whatever they were). The point is not that these propositions *cannot* be understood, it is just that in understanding the ISTA one does not also understand the thoughts to which the attribution is being made.

DSTAs are such that one *cannot* understand them without knowing to which representation truth is being attributed. One cannot understand that 'snow is white' is true without knowing that truth is being attributed to the sentence 'snow is white'. One cannot

⁵ To be explicit, I am not thinking of the attributions as linguistic entities, but as thoughts or propositions. DSTAs and ISTAs are thoughts or proposition that involve attribution of truth to a representation or representations (which themselves might be thoughts or propositions).

understand that the thought that snow is white is true without knowing that truth is being attributed to the thought that snow is white.

I will primarily be concerned with the cognitive value of DSTAs, and in particular of DSTAs that attribute truth to certain kinds of representations, namely *thoughts* or *propositions*. I will call these *propositional DSTAs*. Let me elaborate on what I will call our *intuitive understanding* of propositional DSTAs. What I have said so far, applied to propositional DSTAs, is that one cannot understand a DSTA without knowing to which thought or proposition truth is being attributed. But again something further can be said: *that in understanding propositional DSTAs, knowing to which thought truth is being attributed is grounded in thinking or entertaining that very thought.*⁶

Distinguishing this claim from related but significantly less plausible claims can bring out its content and plausibility. These claims are:

- a. In understanding DSTAs **in general** (and not just propositional DSTAs) knowing to which **representation** truth is being attributed is grounded in thinking or entertaining the thought (for propositional DSTAs) or **the thought expressed by the representation** (for other DSTAs) to which truth is being attributed.
- b. In understanding (propositional) DSTAs, knowing to which propositions truth is being attributed is grounded in thinking or entertaining **with complete understanding** the thought to which truth is being attributed.
- c. In understanding truth attribution that involve **a direct way of specifying propositions**, knowing to which propositions truth is being attributed is grounded in thinking or entertaining the thought to which truth is being attributed.

(a) is to be rejected because one can understand truth attributions to sentences or utterances without knowing which proposition or thought is expressed by the sentences or utterances to which truth is attributed. Then, for example (1) will be such that it can be understood without thinking or entertaining the thought expressed by ‘snow is white’. (b) should be rejected, first, because incomplete understanding is ubiquitous, and second, because on the positive account that I will give here, the cognitive value of DSTAs will be precisely to make the kind of improvement involved in the correction of understanding possible.⁷

(c) should be rejected as well. Suppose that in trying to make the axiomatic character of arithmetic clear to her students, a math professor tells her students that arithmetical truths can be ‘boiled down’ to some basic *truths* called *axioms*. Suppose that she thinks that propositions are the primary bearers of truth. She writes ‘0 is a number’ and ‘For all n such that n is a number, $S(n)$ is a number’, etc., expressing the

⁶ For related discussion, but focusing on propositional attitude attributions, see Tyler Burge’s postscripts to ‘Frege and the Hierarchy’ in Burge 2005 and ‘Belief *De Re*’ in Burge 2007, and Christopher Peacocke’s chapter ‘Representing Thoughts’ in Peacocke 2008.

⁷ I argue this point in detail in my ‘Truth, Inc.’

Peano–Dedekind axioms on the board and labels them, with the labels naming the axioms themselves: ‘Axiom 1’ is a name of the proposition that zero is a number, ‘Axiom 2’ is the name of the proposition that if n is a number, then the successor of n is a number, etc. Now consider

7. Axiom 2 is true
8. That proposition is true [uttered by S in a context where S intends to refer to Axiom 2]

We can allow that there is a sense in which understanding the thought or proposition expressed by a sentence containing these devices of direct reference entails knowing which object the thought is about. But, applied to this case of making reference to a proposition, this sense of knowing which object the thought is about (in this case the thought if n is a number then the successor of n is a number) is consistent with not thinking or entertaining that very thought. Suppose a math professor succeeds in getting the students to understand what an axiom is, but students fail to recognize that the ‘S’ the professor uses on the blackboard expresses the concept of the successor function, and thus fail to understand Axiom 2 (the thought or proposition).⁸ Then the students will understand (7) and (8) without thinking or entertaining the thought that if n is a number, then the successor of n is a number (for related argument see Soames 1999: 48–49).

So it seems that propositional DSTAs involve a mode of reference to a proposition that provides a way of knowing which proposition is in question that is *even more direct than* that provided by demonstrative reference to a proposition or reference with the name of a proposition. Propositional DSTAs exploit a way of knowing which object is in question that exists only for entities like thoughts and propositions, for they are alone amongst kinds of objects in being not only such as to be able to be thought *of*, but also such as to be able to be thought *with*.

This suggests three aspects to our intuitive understanding of DSTAs. Understanding a DSTA requires

- U1. Understanding a directly specifying concept of a thought in making reference to a thought.
- U2. Understanding the thought to which the attribution is being made.
- U3. Understanding the attribution of truth to the thought directly specified.

With this construal of our intuitive understanding of DSTAs in mind, we can ask not only whether this or that view of the concept of truth assigns cognitive value to DSTAs, but in a more precise way, ask whether a view assigns cognitive value to DSTAs intuitively understood.⁹

⁸ I do not claim that the students cannot understand the concept of the successor function, only that in this case they understand what is expressed by (7) and (8) without understanding Axiom 2 because they do not know that ‘S’ expresses a concept of the successor function.

⁹ My account of what is intuitively understood in understanding truth attributions, and my focus on the cognitive value of the concept of truth, overlap to some degree with what Gila Sher has called a “‘core’

III

In this section, I explicate the deflationary and inflationary views of the concept of truth in terms of cognitive value. Space limitations necessitate that the explanation be at a broader strokes level than is ideal even for the programmatic and exploratory level of explication being pursued in this paper. I also explain why neither the deflationary nor the inflationary view can be understood as assigning cognitive value to DSTAs intuitively understood.

Two ideas frame my explication of the deflationary and inflationary views in terms of cognitive value. The first is that that they are helpfully understood as sophistications of more traditional views—the deflationary view of the *redundancy* view and the inflationary view of more traditional metaphysically heavyweight views like the correspondence, coherence, and pragmatist views of the concept of truth. The second is that what the views become more sophisticated about is how to hold to the idea that the concept of truth does (the inflationary view) or does not (the deflationary view) have cognitive value.

I begin with the deflationary view as a sophistication of the *redundancy* view. According to the redundancy view, what is expressed by

1. ‘Snow is white’ is true

and

3. The thought that snow is white is true

is identical to what is expressed by

- 3R. Snow is white

Despite appearances, DSTAs involve neither reference to a representation nor attribution of truth. As Soames put it, ‘[a]ccording to the classical redundancy theory...to say that the proposition that the earth is round is true is not to refer to a proposition and not to describe it as being a certain way’ (Soames 1999: 232). So use of the concept of truth in DSTAs does not allow one to think anything that one could not think already without the concept of truth. In this sense, the concept of truth is redundant.

The deflationary view improves on the redundancy view in two ways. First, as I have presented it, the redundancy view says nothing about ISTAs. The deflationary view does. Second, the redundancy view fails to explain away reference *to* thoughts or propositions (and representations more generally) and thus fails to explain how thoughts that make reference to propositions (like (3)) can have the same cognitive value as a thought that do not (like (3R)); for from thoughts of the first kind it follows that thoughts

unifier” (Sher 2004: 24) for the concept of truth, namely the *Immanence Thesis*: ‘Truth, according to the Immanence Thesis, lies at the juncture of three basic principles of human thought: immanence, transcendence, and normativity’ (Sher 2004: 26; cf. also Sher and Wright 2007: §5). These correspond roughly to my (U1)–(U3). Unfortunately space restrictions prevent a comparison of my view with Sher’s (and Sher and Wright’s).

exist, but this does not follow from thoughts of the second. The deflationary view explains this.

According to the deflationary view, the concept of truth is not redundant. The deflationary view is in agreement with the redundancy view on the following: that were DSTAs to exhaust the logico–linguistic distribution of the concept of truth, then the concept of truth would be without cognitive value (see Field 1994: §5; Horwich 1998: 2ff., 40; Soames 1999: 23; Williams 1999: 547). But the concept of truth is not idle, on the deflationary view, precisely because DSTAs do not exhaust the logico–linguistic distribution of the concept of truth. The value of the concept of truth is to allow the formulation of ISTAs. But what is supposed to make these views deflationary is that the cognitive value of ISTAs is exhausted by the cognitive value of DSTAs, which in turn, are construed as not having any cognitive value, despite the reference to thoughts. The deflationary idea is tie up the loose ends of the redundancy view into a neat little bow, but for a package that still contains nothing of cognitive value inside.

Although this material is familiar, it will be instructive to make things a little bit more explicit, and especially to mark the role of the Fregean criterion in the deflationary view. As a first step, note that an ISTA like

5. Everything Plato believed is true

expresses a universal generalization over thoughts or propositions, to the effect that

5F. For all p , if Plato believed p , then p is true

Second, this universal generalization is supposed to be cognitively equivalent to the infinite conjunction of conditionals

5IT. [If Plato believed the thought that snow white, then the thought that snow is white is true] and [if Plato believed the thought that grass is green, then the thought that grass is green is true] and....

This move from (5F) to (5IT) involves a first application of the Fregean criterion, with the idea being that one cannot, without irrationality, take conflicting cognitive attitudes to (5F) and (5IT). Finally, the instances of the schema (T)

(T) The thought that p is true iff p

are supposed to license substitution in (5IT) yielding (5I), a thought which does not involve the concept of truth:

5I. [If Plato believed the thought that snow white, then snow is white] and [if Plato believed the thought that grass is green, then grass is green] and....

The move from (5IT) to (5I) involves a second application of the Fregean criterion, for the instances of (T) license substitution only because the right– and left– hand sides are cognitively equivalent, i.e. are such that one cannot, without irrationality, take conflicting

cognitive attitudes to (5) and (5I). The instances of (T) express a cognitive equivalence despite the fact that they involve *semantic ascent*—reference to thoughts. Although the reference to thoughts is important because there is no route from (5I) back up to (5F) bypassing the instances of (T) and using standard logical resources (in particular using an objectual understanding of the quantifiers), it functions ultimately as a mechanism of mere abbreviation. This is the result of a third application of the Fregean criterion, one that clarifies the response on the behalf of the deflationary view to the general challenge that ISTAs pose of constituting examples of truth attributions with cognitive value. On the deflationary view ISTAs do not have cognitive value because one cannot without irrationality take conflicting attitudes to, for example, (5) and (5I), but (5I) does not involve the concept of truth. The cognitive value of the concept of truth is to effect this abbreviation of sentences of infinite length that do not involve any expression of the concept of truth into sentences of finite length that do involve the expression of the concept of truth—sentences like (5). So, like the redundancy view, but via a more complex route, the deflationary view also understands the concept of truth to be of nil cognitive value. The concept of truth is not redundant, but it is a device of mere abbreviation.

The three applications of the Fregean criterion mark two direct, and one indirect, targets for the inflationary view. But it will be instructive to work into this view from some more traditional and metaphysically heavyweight accounts of the cognitive value of the concept of truth like the correspondence, coherence, and pragmatist views of the concept of truth.

These views inflate the cognitive value of the concept of truth by making truth a substantial property: of corresponding to the facts, of cohering with a system of beliefs, of what works and is useful. Focusing for example on the correspondence view as representative, the view holds that the concept of truth is not redundant, and the concept of truth is not without cognitive value. The cognitive value of a propositional DSTA like

3. The thought that snow is white is true
is:

3C. The thought that snow is white corresponds to the facts.

On this way of thinking of the concept of truth, the cognitive value of the concept of truth is to provide a way of thinking of a special bit of reality, the correspondence relation. The correspondence view understands cognitive value in terms of the slogan mentioned earlier: TRUTH is a concept of the correspondence relation, so the cognitive value of the concept of truth is to make possible thinking and knowledge of the correspondence relation as TRUTH. Or, put a bit more clearly, the cognitive value of the concept of truth is to make possible thinking and knowledge of the thought that the thought *p* corresponds to the facts as the thought THE THOUGHT THAT *p* IS TRUE.

The inflationary view can be culled from Anil Gupta's (1993) critique of the deflationary view. In outline, the inflationary view explains the cognitive value of the concept of truth not by having the concept of truth invoke a special bit of reality, but by making the concept of truth a special kind of concept for thinking about ordinary bits of reality. Let me explain.

Two key ideas propel Gupta's criticism of the deflationary view.¹⁰ Of primary interest here will be the second idea. The first idea is that the equivalence between an ISTA and an infinite conjunction or disjunction is not synonymy or cognitive equivalence, but rather something weaker. As Gupta puts it, 'a universal statement... does not have the same sense as the conjunction of its instances. The two typically do not even imply the same things; they are equivalent only in a much weaker sense' (Gupta 1993: 289).¹¹ But if the equivalence between an ISTA and an infinite conjunction or disjunction is not that of cognitive equivalence, then, proximately, the move from (5F) to (5IT) and from (5IT) to (5F) is blocked; and more distally, the idea that the concept of truth functions to produce abbreviations of thoughts that do not involve the concept of truth is jeopardized.

Gupta's second key idea concerns the *ideology* of the concept of truth as it is understood on the deflationary view. The ideology of a concept is 'the totality of concepts employed in the definiens of the definition' (Gupta 1993: 298) for the concept. Now if the concept of truth is such that DSTAs are cognitively equivalent to the thoughts to which they are attributed, then understanding DSTA requires being able to think the thought to which truth is attributed, and thus to have the conceptual resources to do so. But this is to attribute an implausibly massive ideology. So, Gupta concludes, DSTAs are not cognitively equivalent to the thoughts to which they attribute truth. This, proximately, blocks the move from (5IT) to (5I), and, again, more distally, jeopardizes the idea that the concept of truth functions to produce abbreviations of thoughts that do not involve the concept of truth.

Gupta concludes, "the meaning and function of 'true' are not connected in the neat way that deflationists supposed" (Gupta 1993: 290). More specifically, Gupta concludes that the deflationary views of the concept of truth 'overlook its most puzzling feature', what Gupta calls 'its remarkable double character': "(i) that an understanding of 'true' requires only a limited range of conceptual resources, and yet (ii) 'true' enables us to talk and think about things that lie far beyond this range" (Gupta 1993: 306). Two versions, a weaker and stronger, of this idea follow from Gupta's key ideas, but it is the stronger version, springing from his second key idea, concerning ideology, that is most relevant. If Gupta's second key idea is correct, it follows that understanding a DSTA cannot require (U2) understanding the thought to which truth is being attributed. For if understanding a DSTA in general requires (U2) understanding the thought to which truth is being attributed, then our understanding of the concept of truth will involve exactly the ideological excess for the concept of truth to which Gupta is objecting. Gupta's idea thus seems to be that in understanding a DSTA without (U2) understanding the thought to

¹⁰ Let me stress that Gupta does not speak primarily about truth attributions to propositions, but truth attributions to sentences. I adapt his view to truth attributions to propositions so as to be able to bring out a radical—indeed too radical—kind of inflationism. The view is more explicit in Collins (2007), whom I discuss in detail in my 'Truth, Inc.'

¹¹ Application of the Fregean criterion explains this, for one can believe that all the instances of a universal generalization are true without believing of all the instances that they are all the instances, and thus without believing the universal generalization (cf. Soames 1999: 247–248). Conversely, as previously noted about ISTAs in general, one may believe a generalization that generalizes over thoughts and attributes truth to them without thinking the thoughts being generalized over, and thus without believing that the instances of the generalization are true.

which truth is being attributed, one comes to be able to think about things that lie beyond what is available for thinking with one's conceptual resources. The cognitive value of the concept of truth lies not in its being a concept of a special thing, for example of the correspondence relation, but in being a special concept that provides cognitive access to ordinary things even in the absence of concepts of those things.¹²

It should be clear then that neither the deflationary nor inflationary views assign cognitive value to DSTAs intuitively understood. Deflationary views may indeed construe DSTAs in accordance with their intuitive understanding, but the concept of truth is of nil cognitive value, serving only to effect abbreviations of thoughts that do not involve the concept of truth. Inflationary views, by contrast, do assign cognitive value to the concept of truth, but as is evident in the discussion above they do so only by failing to construe DSTAs in accordance with their intuitive understanding. In particular, on the inflationary view, understanding a DSTA does not require (U2) understanding the thought to which truth is being attributed.

So explicating the concept of truth in terms of the notion of cognitive value allows one to see that a broad range of views on the concept of truth agree in thinking that DSTAs intuitively understood do not possess cognitive value. But is it correct that DSTAs intuitively understood do not have cognitive value, or is this somehow a blind spot in thinking about the concept of truth?

V

In this section, I defend the idea that DSTAs intuitively understood *do* possess positive cognitive value. The conclusion, in brief, is going to be that the cognitive value of the concept of truth in DSTAs is to combine with a *metarepresentational device of dual reference* to attribute truth to a thought that is simultaneously *thought about* and *thought with* in the course of *critical reflective thinking*.¹³ My argument will proceed by way answering the question: what kinds of thoughts must one be capable of understanding in order to engage in critical reflective thinking? The answer will be: nothing short of thoughts in which the concept of truth combines with a metarepresentational device of dual reference to attribute truth to a thought that is simultaneously thought about and thought with.

What kinds of thoughts must one be capable of understanding in order to engage in critical reflective thinking? I break this question down into four sub-questions. The answers to these sub-questions will constitute an argument for the claim that DSTAs intuitively understood are partially constitutive of the epistemic advance made available in critical reflective thinking.

¹² Gupta characterizes the double character of the concept of truth differently later in his paper, as: '(i) that it enables us to talk about the inexpressible contents of *L*, but (ii) the explanation of its meaning does not appeal to those contents' (Gupta 1993: 307). I think that (i) should read: 'that it enables us to talk and think about things, concepts of which are inexpressible in *L*'.

¹³ My account of the cognitive value of the concept of truth in terms of the constitutive role of DSTAs in critical reflective thinking is intended to be pitched at the same level that Hartry Field (2001) and John Collins (2007) use the phenomenon of *incorporation* in their accounts of 'true' or the concept of truth. Indeed, I argue against the 'incorporation model' (Field 2001: 149) as a competing account of the cognitive value of the concept of truth in my 'Truth, Inc.'

Q1: What is critical reflective thinking?

Reflective thinking, as I will be thinking of it here, is a cognitive resource for epistemically improving one's beliefs under conditions of fixed evidence. Reflective thinking does not provide additional evidence but improved clarity over exactly what the evidence supports and how. I distinguish two kinds of reflection, *simple reflective thinking* and *critical reflective thinking* (Rattan 2008: 241). I will be focusing on critical reflective thinking, but it will be instructive to begin by getting clear on what is involved in simple reflective thinking.

Simple reflective thinking is a matter of one's taking special care in sticking to the normative rules. It is a matter of one's ensuring that one's attitudes are as they ought to be given one's evidence and given the normative rules. Simple reflective thinking is a commonplace. For example, when we check a proof, we are often, perhaps almost always, checking whether we have stuck to the rules for constructing and taking steps in a proof. Was the axiom written down with the quantifiers in the right order? Was the inference rule applied under the right conditions? Has the definition been correctly implemented? In simple reflection, we aim to try to conform our epistemic practice to normative rules that we take as given.

By contrast, critical reflective thinking involves one's scrutinizing *the entire normative relation* that holds between attitudes and evidence—one's attitudes, evidence as well as the normative rules that relate them. In critical reflective thinking, nothing is immune from epistemic evaluation, including the normative rules that one ordinarily takes as given in coming to judgment and belief. A hard question is how the kind of scrutiny involved in critical reflective thinking is possible in application to basic normative rules. The question re-capitulates the basic dilemma for the epistemic evaluation of foundations: to evaluate without begging the evaluative question. But it is not at all clear how one can step back from basic epistemic normative rules to gain an epistemic perspective *on them* that is epistemically independent *of their use* in cognitive thinking.¹⁴ The general problem is that it is not clear against what standard the *basic* normative rules can be held. The idea, to which I now turn, is that such normative rules are evaluated against a norm of truth.

Q2: Suppose that one could not understand truth attributions at all. Would one be able to engage in critical reflective thinking?

¹⁴ Let me briefly mention two kinds of examples. First, suppose one comes to hold a belief on the basis of a deductively valid inference from other beliefs. Critical reflective evaluation of this belief evaluates the entire normative relation between belief and evidence including the normative rules that connect belief and evidence. In this case, epistemic evaluation of the relevant norms is epistemic evaluation of deductive inference. But this is the problem of the justification of deduction; and a significant aspect of that problem is to explain just what form a non-question begging justification of deduction is to take. Second, suppose that one reasons with the analysis of some term or concept *C* together with specific information about an object *o* to a concluding belief that *o* is *C* (for example *o* could be a function, and *C* the concept of continuity). Suppose, however, that it is highly counterintuitive that *o* is *C*, and the specific information about *o* is not in doubt. Then critical reflective evaluation of the belief will involve epistemic evaluation of the analysis of *C*. But this comes up against the paradox of analysis. For how could we come to correct the analysis of *C*—would that not make the *analysis* of *C* *informative*, and if so, how is that possible?

In other work (Rattan 2008) I argue, against some sceptical suggestions of Paul Horwich's (2006) concerning the cognitive value of the concept of truth, that critical reflective thinking is the aspect of our epistemic practice in which the value of truth is manifest.¹⁵ Horwich takes it as an intuitively plausible starting point that our epistemic practices are justified and explained by a concern for the value of truth (2006: §7). But against this he argues for and accepts the conclusion that our epistemic practices are neither justified nor explained by a concern for the value of truth. In Rattan (2008) I argued that the value of truth is manifest in critical reflective thinking by being partially constitutive of the kind of epistemic advance made in critical reflective thinking. The exclusionary point of this thesis is not to deny that other epistemic resources, for example, perception or memory, are veridical by their nature. But the veridicality of perception and memory do not make the value of truth manifest. An understanding of truth attributions plays no constitutive role in the epistemic advances made (or made possible) by perception and memory. But the epistemic advance made in critical reflective thinking is partially constituted by an understanding of truth attributions. The epistemic advance in critical reflective thinking is one in which normative rules that connect belief and evidence are themselves subject to a standard of epistemic evaluation. But what standard is that? Returning to our intuitively plausible starting point: it is the standard of truth. In engaging in critical reflective thinking, it is manifest to the thinker that she assigns cognitive value to the concept of truth because in engaging in critical reflective thinking she is engaging in a form of epistemic evaluation that requires her to grasp the concept of truth. In critical reflective thinking one is required to understand truth attributions in order to go beyond the question of whether one's belief is in accordance with basic normative rules to the question of whether one's belief is true.

A contrast with simple reflection helps to make this clearer. Consider a relativist who does not think that there is a norm of truth governing her attitudes. The relativist could nevertheless allow that her attitudes are governed by norms that are contingent, historical, social, or local. And with these norms given, she can engage in simple reflection to conform her attitudes to these norms. This is consistent with her relativism, for there is nothing in simple reflection that goes beyond the idea of correctness relative to a standard, for simple reflection checks only whether attitudes are as they ought to be *given* basic normative rules that are supposed to normatively relate attitudes to evidence. So simple reflection is consistent with relativism, a view in which there is no governing norm of truth. So simple reflection cannot be where in our epistemic practice the cognitive value of the concept of truth is manifest. But because basic norms are themselves under critical scrutiny in critical reflective thinking, critical reflective thinking cannot be understood as aiming for correctness relative to norms; it aims for correctness unrelativized. The intuitive starting point is that this norm is truth. When we are engaged in inquiry that asks not only after conclusions, or their backing, but also the cogency of that backing—issues about methodology, reasoning, and the rationale for basic inferences and belief forming methods—we are engaged in critical reflective thinking, and such thinking cannot be understood as sticking to the rules. It is thinking governed by the norm of truth.

¹⁵ I think now a better term would have been 'immanent'. Cf. Sher 2004.

So critical reflective thinking concerning one's belief that p is not possible without being able to understand truth attributions. So someone who could not understand truth attributions could not engage in critical reflective reasoning. But this does not yet show that understanding DSTAs in particular is partly constitutive of critical reflective thinking. But further consideration of the nature critical reflective thinking brings this out.

Q3: Suppose that one could understand only ISTAs and not DSTAs. Would one be able to engage in critical reflective thinking?

The nature of reflective thinking is shot through with metarepresentation. Both simple and critical reflective thinking involve a conceptualization of one's attitudes, their evidence, and the normative rules that relate attitudes and evidence. There is metarepresentation of attitudes understood as cognitive relations to thoughts; of thoughts; of evidence; and of the normative rules that normatively relate belief and evidence. What may have appeared a wide gulf between the cognitive value of the concept of truth and the nature of critical reflective thinking begins to appear much narrower once the metarepresentational aspects of critical reflective thinking are made explicit. So critical reflective thinking and truth attributions both make use of metarepresentational thinking. But is critical reflective thinking partially constituted of metarepresentational thinking of a certain sort? Even if understanding truth attributions is required, is understanding of DSTAs *in particular* required?

Critical reflective thinking expresses a thinker's considered point of view. A thinker's considered point of view might itself be newly occasioned through the employment of the cognitive resource of critical reflective thinking. That is, critical reflective thinking is such as to be able to produce *a change in one's attitudes*, for example, a change in which one moves from what is considered, on reflection, to be a rationally suboptimal epistemic state, say one of believing that p , to what is considered, upon reflection, to be the epistemically improved state, say of believing that q . But one cannot engage in such a revision of belief without being able to identify which beliefs one has. It is not good enough to think of the thought that one believes descriptively, for example as the thought Socrates expresses with his first declarative sentence in *Meno*. This is not a way of thinking of one's belief that allows it to be the object of revision in critical reflective thinking that aims to express one's considered point of view.

Consider also the point at which one considers the simple reflective question of whether one's beliefs are in accordance with some norm, like a norm of consistency for belief. Again it will not be enough to think of one's thoughts specified descriptively, for example as the thought Socrates expresses with the first declarative sentence of *Meno*. Evaluating whether that belief is an accord with the evidence and the normative rules that relate belief and evidence requires a more direct specification of the thought that is the object of that belief. This is because the normative rules that relate belief and evidence will specify thoughts, or kinds of thoughts, directly and not indirectly and will tell us, for example with consistency, that revision is called for if one has the belief that snow is white and snow is not white, or, more generally, if one has belief of the form p and not- p . Evaluation of attitudes with respect to the norms will require what we might call *directly*

specified belief attributions to mesh with the statement of the normative rules that relate belief with evidence.

The same point applies for critical reflective thinking and the evaluation of norms themselves. The norms will specify beliefs or belief types directly, and unless one can identify the norms, one will not be in a position to evaluate them.

If that is right, then understanding of DSTAs in particular is partially constitutive of critical reflective thinking, for without the use of DSTAs in particular neither will one be able to identify the attitudes that require revision, nor will one be able to revise in accordance with the normative rules, nor will one be able to identify the normative rules themselves.

Q4: Suppose that in order to engage in critical reflective thinking, one must be able to understand DSTAs. Must they be construed in accordance with their intuitive understanding, or can they be understood in some other, less intuitive, but perhaps theoretically supported, way?

Critical reflective thinking is a matter of evaluating the entire normative relation between belief and evidence. In the light of a critical reflective evaluation of one's beliefs, one may modify one's belief because one's belief is not as it should be given what are on reflection correct normative rules that relate belief and evidence, or because one's beliefs took as given normative rules that upon reflection are mistaken. In either case, in order for one to engage in critical reflective thinking, one needs to be able not only to think *about* the thought that is the object of one's belief, but also to be able to think the thought itself because it is the content of *one's own* belief. So one could not engage in critical reflective thinking without being able both to think about and with the same thought. Given the earlier result that critical reflective thinking requires an understanding of truth attributions, and in particular DSTAs, it follows that one cannot engage in critical reflective thinking without understanding DSTAs in accordance with their intuitive understanding.

VI

I conclude with three briefly presented lessons, one at each of the levels of explication pursued in this paper. I think that each lesson fully elucidated would involve its own paper length treatment, so what I say in this conclusion will be of necessity incomplete.

At the programmatic and exploratory level, I pressed the notion of cognitive value into the service of explicating the concept of truth by construing a broad range of views of the concept of truth in terms of cognitive value, in finding in them a common blind spot to the positive cognitive value of DSTAs, and in providing my own account of the cognitive value of DSTAs in terms of the epistemic advance made in critical reflective thinking. But what bears special emphasis is that whereas the explanation of how the deflationary and inflationary views account for cognitive value makes use of the Fregean criterion, the explanation of cognitive value in my account does not. My account does not explain the cognitive value of DSTAs by showing how DSTAs differ in cognitive value from the thoughts to which they attribute truth. Rather my account of the cognitive value

of truth relies on the fact that in critical reflective thinking *one and the same thought* is both thought about and thought with. What engenders cognitive value is not a cognitive difference or a difference in dubitability, but rather a cognitive sameness, namely that of a thought one is thinking with and a thought one is thinking with. This is an application of the notion of cognitive value that is governed neither by the slogan explaining the cognitive value of a concept, nor by the by Fregean criterion.¹⁶

At the dialectical level, my view stakes out a position between the deflationary view and the inflationary view. Like both the inflationary and deflationary views, my view eschews appeal to a metaphysically substantial entity, like the correspondence relation, in explaining the positive cognitive value of DSTAs. Like the inflationary view, my view assigns cognitive value to DSTAs, but not by allowing thinkers to transcend their conceptual resources—not by allowing thinkers to access with DSTAs what they cannot access with their conceptual resources. My view rejects this idea because it is inconsistent with construing DSTAs in accordance with their intuitive understanding, and in particular with the idea that understanding a DSTA requires (U2) understanding the thought to which truth is being attributed. Rather, my view assigns cognitive value to DSTAs for their constitutive role in critical reflective thinking, thinking that exploits for cognitive value that one and the same thought is both being thought about and thought with. I think but have not shown here that the cognitive value of truth lies in allowing thinkers not to transcend, but to *attain mastery*, over their conceptual resources.¹⁷

At the substantial level, the cognitive value of the concept of truth, in particular in DSTAs, is to combine with a metarepresentational device of dual reference to attribute truth to a thought that is simultaneously thought about and thought with in the course of critical reflective thinking. A fuller defence of this view would require a fuller and comparison of this view against the inflationary view, and in particular against the view in which the cognitive value of the concept of truth derives from the constitutive role that it plays in *incorporation*.¹⁸

References

Burge, T. (1993). Content Preservation. *Philosophical Review*, 102: 457-488

Burge, T. (2005). *Truth, Thought, Reason: Essays on Frege*. New York: Oxford University Press

Burge, T. (2007). *Foundations of Mind: Philosophical Essays Volume 2*. New York: Oxford University Press

¹⁶ This aspect of cognitive value or significance is discussed more fully in Dickie and Rattan (forthcoming: §3) under the head of ‘the consolidating explanatory role of sense’. Related are the notions of content preservation in Burge 1993 and coordination in Fine 2007.

¹⁷ For some discussion of this, see my manuscript ‘Intellect and Concept’.

¹⁸ See my ‘Truth, Inc.’ for further details.

- Collins, J. (2007). Declarative Thought, Deflationism and Metarepresentation. In D. Griemann and G. Siegart (eds.), *Truth and Speech Acts: Studies in the Philosophy of Language* (157–177). London: Routledge
- Dickie, I. and Rattan, G. (forthcoming). Sense, Communication, and Rational Engagement. *dialectica*
- Evans, G. (1982). *Varieties of Reference*. Oxford: Oxford University Press.
- Field, H. (1994/2001). Deflationist Views of Meaning and Content. Reprinted in H. Field, *Truth and the Absence of Fact* (104–140). New York: Oxford University Press
- Field, H. (2001). Postscript to ‘Deflationist Views of Meaning and Content’. In H. Field, *Truth and the Absence of Fact* (141–156). New York: Oxford University Press
- Fine, K. (2007), *Semantic Relationism*. Oxford: Blackwell
- Frege, G. (1892/1997). On Sense and Reference. Reprinted in M. Beaney (ed.), *The Frege Reader*. Oxford: Blackwell
- Goldberg, S. (2008). Must Differences in Cognitive Value be Transparent? *Erkenntnis*, 69: 165–187
- Gupta, A. (1993/1999). A Critique of Deflationism. Reprinted in S. Blackburn and K. Simmons (eds.), *Truth*. Oxford: Oxford University Press
- Horwich, P. (1998). *Truth*. Oxford: Blackwell
- Horwich, P. (2006). The Value of Truth. *Noûs*, 40: 347–360
- Kimbrough, S. (1998). Anti-Individualism and Fregeanism, *Philosophical Quarterly*, 48: 470–482
- Peacocke, C. (1992). *A Study of Concepts*. Cambridge: MIT Press
- Peacocke, C. (2008). *Truly Understood*. New York: Oxford University Press
- Quine, W.V.O. (1970). *Philosophy of Logic*. Englewood Cliffs: Prentice Hall
- Rattan, G. (2008). On the Value and Nature of Truth. *Journal of Philosophical Research*, 33: 235–251
- Rattan, G. (forthcoming). Intellect and Concept. In B. Armour-Garb, D. Patterson, and J. Woodbridge (eds.), *The Baltic International Yearbook of Cognition, Logic and Communication*, 4

Rattan, G. (manuscript). Truth, Inc.

Sher, G. (2004). In Search of a Substantive Theory of Truth. *Journal of Philosophy*, 101: 5–36

Sher, G. and Wright, C. D. (2007). Truth as a Normative Modality of Cognitive Acts. In D. Griemann and G. Siegart (eds.), *Truth and Speech Acts: Studies in the Philosophy of Language* (525–574). London: Routledge

Soames, S. (1999). *Understanding Truth*. New York: Oxford University Press

Williams, M. (1999). Meaning and Deflationary Truth. *Journal of Philosophy*, 96: 545–564