



HAL
open science

The False Modesty of the identity theory of truth

Pascal Engel

► **To cite this version:**

Pascal Engel. The False Modesty of the identity theory of truth. *International Journal of Philosophical Studies*, 2001, 9 (4), pp.441-459. <ijn_00000317>

HAL Id: ijn_00000317

https://hal.science/ijn_00000317v1

Submitted on 26 Jan 2003

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

THE FALSE MODESTY OF THE IDENTITY THEORY OF TRUTH

University Pascal Engel
of Paris-Sorbonne

International Journal of Philosophical Studies, 9, 4, 2001, 441-458*

Abstract : The identity theory of truth, according to which true thoughts are identical with facts, is very hard to formulate. It oscillates between substantive versions, which are implausible, and a merely truistic version, which is difficult to distinguish from deflationism about truth. This tension is present in the form of identity theory that one can attribute to McDowell from his views on perception, and in the conception defended by Hornsby under that name.

Several writers have recently argued that to the familiar list of theories of truth one should add the “identity theory of truth”, the view that the truth of a judgement consists in its *identity* with a fact. It played an important role in Bradley and in the early writings of Moore and Russell (Cartwright 1987, Candlish 1989, Baldwin 1991). It seems also to play a role in McDowell’s *Mind and World* (Dodd 1995) and Hornsby (1997) has defended a view under that name. There are various reasons why one might want to develop a view of this kind, but in the context of the present day debate between “robust” and “minimalist” conceptions of truth [\[1\]](#), one line of reasoning apt to motivate the adoption of an identity theory of truth might be the following. (a) It is a truism, indeed one of those on which a “deflationist” conception of truth insists, that “it is true that *p*” and “it is a fact that *p*” are equivalent ways of saying the same thing. (b) If one wanted to read more into this truism, one would have to understand truth as a form of correspondence between *p* and some worldly item, namely a fact. But a correspondence theory is hopeless. (c) This, however, does not necessarily lead us back to a conception of truth which would say that the meaning of “true” is exhausted by the harmless truism and by “ ‘P’ is true if and only if P”. For it would still be open to us to say that truth consists in a substantive relation, which is not that of correspondence, but that of identity: a true thought *is identical with* a fact (in italics, in order to distinguish this claim from the truism). This kind of reasoning seems to underlie, at least in part, Hornsby’s proposal, when she says that “the identity theory is worth considering to the extent to which correspondence theories are worth avoiding” (1997: 6), and that “to the extent that the minimal theorist wants to convey a deflationary message about truth, which is not already conveyed in the identity theorist opposition to correspondence, the message has to be resisted” (*ibid*, 16).[\[2\]](#) This suggests that we could steer away both from a deflationary conception of truth *and* from a correspondence conception. An identity

theory of truth would acknowledge the extreme proximity of our talk of truths and of our talk of facts, *without* falling back into the view that “fact” is a mere rephrasing of “true thought”, since it insists that truth requires the existence of a relation between thought and reality. What would such a theory look like? And can it have it both ways? I shall try to show here (i) that there is no single view which can be called “the identity theory of truth”, and that no of the interpretation that we can give of it is promising, (ii) that the versions which one might attribute to McDowell and to Hornsby suffer from similar problems.

1. A false start: Frege and the identity theory

As Baldwin and Hornsby have reminded us, an identity theory of truth (henceforth IT) may be formulated from a famous passage in Frege’s “The Thought” where he discusses the correspondence theory:

“A correspondence, ..., can only be perfect if the corresponding things coincide and are, therefore, not distinct things at all. It is said to be possible to establish the authenticity of a bank note by comparing it stereoscopically with an authentic one. But it would be ridiculous to try to compare a gold piece with a twenty mark piece stereoscopically. It would only be possible to compare an idea with the thing if the thing were an idea too. And then, if the first did correspond perfectly with the second, they would coincide. But this is not at all what is wanted when truth is defined as correspondence of an idea with something real. For it is absolutely essential that the reality be distinct from the idea. But then there can be no complete correspondence, no complete truth. So nothing at all would be true; for what is only half true is untrue.” (Frege 1967: 18-19)

Here Frege seems to say that if truth consisted in a correspondence between a thought and reality, the correspondence would have to be perfect - it would have to be an *identity* – which is absurd, since the relation of correspondence implies that two different things have to correspond to each other. But the passage can also be read as a *reductio* of the correspondence view *on behalf* of the relation of identity between thought and items in the world: if there were no identity there would be no truth. Now Frege also says, in the sequel of this passage, that truth is *undefinable*, and, together with his rejection of truth as correspondence, his argument for this claim is his famous “regress”: of any purported definition of truth we would have to say that it is true, and that it is true that it is true that this definition is true, etc.^[3] And he also says, at different places, that there is no more to say about truth than that *it is true that p* is equivalent to the assertion of *p* itself, and that facts are simply true thoughts, but in a trivial sense, and this has often been seen as an anticipation of a minimalist conception.^[4] So it is quite dubious that Frege would have been prepared to go further than step (b) of the reasoning above. Moreover, if he had wanted to identify

true thoughts with facts in the identity sense, he would have blurred his famous distinction between sense and reference, for thoughts are senses, and presumably facts, if they belonged to the world, would be located within the world of references: if true thoughts were facts they would be constituted both by entities of the two realms (Baldwin 1991: 43). So Frege is probably not an identity theorist of truth, and if he was, his IT would oscillate between the harmless truism that a true thought is a fact, and the claim (b) that it is more than a truism.

2. The identity theory of facts

If an IT goes beyond the truism, we have to say more about the notion of fact, and more about the notion of thought. But if we want also to take into account the notion of identity, we have to remember that the identity “A true thought is a fact” can be read both from right to left, nudging thoughts towards the world, and from left to right, nudging the world towards thought.^[5] There are versions of IT which go in one way, and versions which go in the other.

Going in the first direction are the versions of IT which rely on a substantive notion of fact, as distinct entities in the world. As Baldwin has shown, Moore and Russell, in their early period, had a view of this kind. They held that the contents of true thoughts or propositions are facts, and that their constituents, concepts, are real things in the world. They identified the truth bearers with the truth makers (Candlish 1999: 234). Famously Frege balked at this, when Russell told him that Mont Blanc itself, with its rocks and its snowfields, is a real constituent of the proposition that Mt Blanc is more than 4000 meters high.^[6] Later Russell developed a view of facts as independent entities from propositions, but without clearly distancing himself from the view that the structure of facts is distinct from the structure of propositions.^[7] But let us not enter here Russell exegesis, and let us call a theory, whoever has held it, which would identify true propositions with facts, conceived as worldly items, an *identity theory of facts* (ITF). It will have to say something substantial both about facts and about propositions, in order to identify them.

It is difficult to figure out what an ITF amounts to, since the usual views of truth which are based on a substantial notion of fact are not identity theories of facts, but *correspondence* theories of facts. They should spell out independently what a facts are, and proceed from there to explain how the truth making relation can be understood in terms of these entities. It is clear that, on such views, the truth makers are not to be identified with the truth bearers. It is also clear that for such theories invoking facts does not amount to a truistic rereading of the notion of true proposition. For instance one of the leading proponents of such theories, David Armstrong (1997: 19) emphatically insists that facts are not the “tautological

accusatives” of true statements or true propositions. The reason is simple. For a tautological account of facts, there are as many facts as there are truths, and a one-one relation between facts and truth bearers. But on a correspondence theory of facts the truth making relation is one-many, or many-one. To take simple examples, if *p or q* (inclusive *or*) is true, this truth has two truth makers, *p* and *q*. Or for a true existential sentence saying that there is at least a black swan, there are as many truth makers as there are black swans. Conversely, to one truth maker correspond many truths. For instance, if it is true that either *p or q* is true, then the truth maker for *p* is also a truth maker for the disjunctive truth, and for innumerable many other truths (*ibid.*129-130). In other words, facts as truth makers are *not* true propositions. This prevents a correspondence theory of facts from “falling into the gravitational field of a redundancy theory, to their mutual confusion” (*ibid.* 128). But it also prevents a correspondence theory of facts from falling into the gravitational field of an identity theory of facts, if such there be, since the identity of propositions and facts implies that there is a one-one relation between them. Now if this is so, the ITF itself runs the risk of falling into the gravitational field of the tautological accusative theory. So how can we hope to extract from it a substantive theory of facts? In so far as we conceive ITF on the model of a correspondence theory of facts as truth makers, then, it is difficult to resist the thought voiced by Baldwin, that “the identity theory is the result of adding the unnecessary insistence that truth requires a relationship between thought and the world.” (Baldwin 1991: 50)

There is a further difficulty in trying to specify ITF: if a true thought is a fact, how can a *false* thought be identical with a fact? Will there be facts which are objective falsehoods? Or should we say instead that false thoughts are not really thoughts, and that only true thoughts are? Truth and falsity seem to be contingent properties of thought contents: these can be true or can be false. To take up Wittgenstein’s terminology his in his *Notebooks*, thoughts or propositions have two “poles”, truth and falsity, and can instantiate whichever property.^[8] Hence if they are identical with facts, they seem to be identical with merely *possible* facts, and not necessarily with *actual* facts. On this view, facts can obtain or not, and if they don’t they remain only possible. But there is another notion of fact, according to which facts *cannot fail* to obtain, or are *essentially* facts. So ITF can be read in two ways, both for truth and falsity :

- (ITF a) (i) The thought that *p* is true = the fact that *p* contingently obtains
(ii) The thought that *p* is false = the merely possible fact that

- (ITFb) (i) The thought that *p* is true = the essentially obtaining fact that *p*

(ii) The thought that p is false ¹ the essentially obtaining fact that p

According to (ITF a), facts themselves have two poles: <obtaining, not obtaining> (they are *bipolar* (Dokic 1998)). In possible world terminology, a true thought could be true at another possible world if it were identical with a possible fact, but truth in the actual world is defined as the actualisation of a possible fact.^[9] According to (ITF b), facts have only one pole: <obtaining>. Facts which are not actual are *not* facts, but mere “states of affairs” or “virtual” facts. Truth is identity with what is essentially or necessarily a fact, and could not be a fact in other possible worlds. Hence falsity is simply the absence of fact, non facthood. But we might go further and allow facthood also for false propositions. So there is a reading of (ITF b) on which false thoughts are identical with *negative facts*. (ii) is replaced by:

(ITF b) (iii) The thought that p is false = the essential fact that not p

Famously Russell held such a view for some time, until he was too much worried by this consequence of his early identity theory of truth.^[10]

ITF a and ITF b (i)-(ii) do not only rely on two different notions of fact, but also on two different notions of proposition or thought. For ITF a , a single entity, a thought or a proposition, can be true or false, hence identical to an actual or a possible fact. If I think, for instance, that spring has begun, my thought is one thing, and its being true or false are other things. Same thought content, different realisations. This is why ITF a is a bipolar theory of facts. For ITF b (i)-(ii), on the contrary, when I think truly that spring has begun, my thought is a fact, the essential fact that spring has begun, but when I think falsely that spring has begun, my thought is not a fact, since only the fact that spring has begun obtains. Hence it is not a thought, if the identity theory is correct, but a mere representation. It does not have the same content when it is identical to a fact and when it is not. This why ITF b (i)-(ii) is a unipolar theory of facts: only true propositions can be facts.^[11] Only when negative facts are introduced with ITF b (i)-(iii) can one reinstate the bipolarity of facts, but also one has to introduce a bipolarity of propositions or thought as well.

Each view, however, creates damaging problems for an ITF. For ITF a implies that the content of a thought single thought can be identical *either* with an actual fact *or* with a possible fact, hence that the content of a thought is different from what makes it true, since actual and possible facts have different modal properties (having different modal properties presumably implies having different properties). So on ITF a , truth bearers cannot be identical with truth makers. ITF b (i)-

(ii) has the same problem, since it implies that a false content cannot be a real thought, hence cannot be identical with a fact. Only ITF *b(i)-(iii)* does reintroduce the appropriate identity of thought and fact, but at the price of admitting these strange entities, essential negative facts. This is not the end of the matter, for Frege's regress also lurks here. On either version of ITF, the property of truth is replaced by the obtaining relation, which is supposed to be the explanation of the identity of truths and facts. But what, on this view, makes a truth true? Not simply a fact, but *the fact that this fact obtains*, i.e a second-order fact that a first-order fact obtains. Now this second-order fact can be read, according to ITF*a*, as capable of obtaining or not, or, according to ITF*b*, as essentially obtaining. Suppose we take the first alternative. Its obtaining or not will have to be explained by another, third-order, fact, and the regress begins. We have no other means of saying that the fact that the fact that *p* obtains is true. Now suppose that we take the second alternative, according to which a fact cannot fail to obtain. Then the regress is blocked, but we have no other way of explaining the obtaining relation than by saying that it is a primitive, undefinable property. This is actually quite close to Frege's reasoning.

An ITF, therefore does not look very promising. A deflationist would probably say that this confirms his claim that there is no way to extract more from the notion of fact than the truism that facts are merely "tautological accusatives" of true propositions. A correspondence theorist would say that it confirms his attempt to develop a theory of facts as truth makers independent from true propositions.

3. The identity theory of thought

The identity theory of facts nudges thought towards the world. But we might alternatively nudge the world towards thought, and read the identity of thought and fact from right to left. Instead of an identity theory of facts, we would have a identity theory of thought (ITT). A natural way of doing this would be to embrace a version of idealism. If thoughts are representations in the mind, the identity of the world with thought would be secured by a form of phenomenalism, for instance. Or if truth bearers are sentences, an identity theory might take the form of the view that the world is but a collection of sentences, or a text, possibly in some Mallarmean or Derridean spirit. This is not very attractive. It seems more reasonable to understand thoughts as thought *contents*, or as *thinkables* (the important of this term will emerge later). Facts, or reality would be what we are able to think.

Actually this idealist theme lurked in Russell's and Moore's early writings, for when they wanted to identify the world with true propositions and with concepts, their insistence on the reality of these entities could just as well be interpreted as an

identification of the things in the world with concepts and thoughts. [\[12\]](#) But this is hardly surprising since they drew part of their inspiration from Bradley.

One finds in Bradley a reasoning which is quite similar to Frege's: if truth were not the same thing as reality, there would be a difference between the two, but then truth would be defective, hence truth cannot be anything else than reality. [\[13\]](#) But Bradley does not conceive reality as a pluralist system of facts; reality is a monistic coherent system. Hence a judgement is true only if its content belongs to a coherent system of reality, the Absolute. The coincidence, however, between an isolated judgement and reality is never perfect. No truth is ever a perfect truth, except when it coheres with others to become the Absolute. This is why Bradley says that a judgement is always "conditional" and that there are degrees of truth, since the perfect identity between thought and reality "can never be stated". [\[14\]](#) We find here the same pattern as that found in Frege's, Moore's, and Russell's early feeling that truth is an ineffable property which attaches mysteriously to judgements, but within a monist setting. We have an identity theory of thought, but it is unstatable. Since reality as well as the thought which, ideally, would be identical to it, is a coherent whole, Bradley's ITT collapses into a coherence theory of truth. The commentators who insist that he was neither a correspondence nor a coherence theorist are right of course, since they have shown convincingly that he was an identity theorist. But certainly the identity theory falls here into the gravitational field of the coherence theory, just as an identity theory of facts falls into the gravitational field of the correspondence theory. What prevents it from falling into the gravitational field of the deflationary theory is Bradley's doctrine of degrees of truth, but when the Absolute is ideally reached, the world disappears into thought, and conversely.

I have examined several versions of what purports to be an identity theory: two (may be three) identity theories of facts, and the identity theory of thought. Neither of them seem to be susceptible to be formulated satisfactorily: either they lead to implausible conceptions of thoughts and facts (ITF), or they (the ITT) lead us back to truth as coherence. Neither of them seems to fit the bill that was suggested by the reasoning (a)-(c) above, of steering away both from a deflationary conception which would reduce facts to tautological accusatives of true propositions and from suspect substantive conceptions of truth. According to minimalism, there is no harm in saying that that truth is correspondence with facts or reality, since '*P*' is true if and only if '*P*' correspond to the facts is just another way of saying what the schema: '*P*' is true if and only if *P*. This "correspondence platitude" is innocuous. [\[15\]](#) But we could just as well say that there is an *identity platitude* :

'P' is true if and only if 'P' is identical with a fact

saying the same thing as the familiar schema, and equally innocuous. We could call this the *deflationary identity theory of truth*. But it not clear that it deserves the name of a distinctive identity theory, just as the correspondence platitude does not deserve the name of a correspondence theory, if we read it in the deflationary way.

It is not clear, then that an identity theory can have an autonomous status from minimalist theories and from substantive theories, and that it can constitute a genuine alternative to them. An IT could have a claim to autonomy if it could be shown that the deflationary, the correspondence and the coherence theories of truth fall into its gravitational field of the IT. But we have seen that on the contrary the various possible versions of IT fall into their gravitational field. So either it has no autonomous status, and it is a mere rephrasing of a deflationist theory together or it is dubiously coherent. Bradley quite lucidly remarks:

“I must venture to doubt whether ...truth, if that stands for the work of the intellect, is ever precisely identical with fact... Such an idea might be senseless, such a thought might contradict itself, but it serves to give voice to an obstinate instinct.”[\[16\]](#)

But if it is not more than an instinct, a deflationary view of truth can help us to resist it.

4. McDowell's identity theory of perceptual thought

In *Mind and World* (1994: 27), John McDowell writes:

“There is no ontological gap between the sort of thing one can mean, or generally the sort of thing one can think, and the sort of thing that can be the case. When one thinks truly, what one thinks *is* the case. So since the world is everything that is the case (as [Wittgenstein] once wrote) there is no gap between thought, as such and the world.”

This seems to imply a commitment to an identity theory of truth.[\[17\]](#) The reference to Wittgenstein also suggests a Tractarian metaphysical picture, according to which the world is the totality of facts, and truth is identical with the world or reality. But McDowell immediately adds:

“But to say that there is no gap between thought, as such, and the world is to dress up a truism in a high flown language. All the point comes to is that one can think, for instance, that *spring has begun* and the very thing, *that spring has begun*, can be the case. That is a truism, and it cannot embody something metaphysically contentious, like slighting the independence of reality.”(*ibid.*)

The last remark is intended as a disclaimer against any idealist interpretation of his view, assimilating reality to representation. For McDowell, truth bearers are thoughts

in the sense of what *can* be thought, what he calls *thinkables* : not acts of thoughts, but possible *objects* of thought (p.28). True thinkables are facts. If it is an identity theory, it seems to be closer to an ITF . But the second quotation also contains another disclaimer: McDowell denies that what he is proposing is more than a mere truism, the one from which we started, which takes facts are mere tautological accusatives of true propositions. Hence McDowell would thus seem not to be proposing an identity theory in any substantial sense.

This should not be a surprise to McDowell's readers, since he has long advocated a form of deflationist or minimalist conception of truth, for instance in "Anti-realism and the epistemology of understanding":

"There is a truistic connection between the notion of a content of an assertion and a familiar notion of truth (one whose significance we might think of as fully fixed by this connection): the connection guarantees, as the merest platitude, that a correct specification of what can be asserted, by the assertoric utterance of a sentence, cannot but be a specification of a condition under which the sentence is true."(1981: 229, 1999: 319)

The minimalist view of truth is also related to his claim that a theory of *meaning* can only be "modest": if there is no more to truth than the platitude expressed by the disquotational schema " 'P' is true if and only if P", there is no way by which we could account for meaning in terms of substantive truth conditions, assertability conditions, or in other terms:

"The recoil [from realism] has nothing to do with the rejection of the truth-conditional conception of meaning, properly understood. That conception has no need to camouflage the fact that truth conditions are necessarily given by us, in a language that we understand. when we say "'diamonds are hard' is true if and only if diamonds are hard", we are just as much involved on the right hand side as the reflections on rule following tell us we are. There is a standing temptation to miss this obvious truth, and to suppose that the right hand side somehow presents us with a possible fact, pictures as unconceptualized configuration of things in themselves. But we can find the connection between meaning and truth illuminating without succumbing to that temptation." (1998: 255)

The realist temptation alluded to here is as much present in a correspondence theory of facts as it is present in the Myth of the Given that McDowell attempts to expel from our thinking in *Mind and World*: to conceive of truth as a relation to a reality external to our thought is just to succumb to a mythical transcendental realism.

So it is dubious, *prima facie*, that McDowell is proposing an identity theory. But this first impression is wrong. When McDowell introduces the "truism" that truth thinkables are facts, it is in the context of a particular theory of *perception*:

"In a particular experience in which one is not misled, what one takes in is *that things are thus and so*. *That things are thus and so* is the content of an experience[...] But

that things are thus and so is also, if one is not misled, an aspect of the layout of the world: it is how things are (1994: 26)

That things are thus and so is the conceptual content of an experience, but if the subject of the experience is not misled, that very same thing, *that things are thus and so*, is also a perceptible fact, an aspect of the real world.(*ibid*)

Here we do not have a proposal about thought in general, but about perceptual thought, which is said to be conceptual and to have the same structure as that of judgements or propositions. Let us call this an *identity theory of perceptual thought* . It does not lead directly to an identity theory of truth for thought contents *in general*, but it directly leads to an identity theory of truth for perceptual contents: if the content of perception is the same as the content of judgement, and if what is perceived is a fact, then the content of perception is a fact. But, on reflection, this can extend also to judgements which are not of perception, for if they have a propositional structure in the same sense as that in which perceptual contents have a propositional structure, then they too are, when they are true, identical with facts. This is not completely obvious for it might be that non perceptual judgements have extra features, not present in perception, which do not make them “about” the world as directly as perception does. But in so far as they also involve “thinkables”, the implication that for these too an identity theory of truth holds is difficult to avoid. So, after all, there are strong grounds to think that McDowell *does* defend an identity theory of truth.

So we can attribute to McDowell the three following theses:

IT (*identity theory of truth*) A true thought is a fact

ITP (*identity theory of perception*) A true perceptual thought is a fact [\[18\]](#)

CP (*conceptual character of perception*) Perceptual content is conceptual

The problem is to understand the relationships between these theses. One could alternatively understand his argument, in lecture II of *Mind and Word*, as inferring CP from IT and from ITP , or as inferring IT from CP and ITP. But the former inference would beg the question, since IT has to be understood as a claim about *propositional* thought, as made up of *conceptual* contents, in order to derive from it, together with ITP, the conceptual nature of perceptual thought CP. If some kinds of thoughts are not propositional, neither ITP nor CP can hold. And this is just the point of McDowell’s attack, in lecture III, on the notion of non-conceptual content. In addition, there are as well thoughts the content of which is not perceptual, so the appropriateness of an IT for them, as we just saw, is not obvious. Therefore McDowell needs some independent motivation for CP, in order to derive it from IT and ITP. On the other hand, if we take his argument in the reverse direction, as

moving from CP and ITP to IT, we need to argue for CP first, and also to be able to generalise what is supposed to be true of perceptual thoughts to other kinds of thoughts, and we have just seen that it is not evident. So the identity theory cannot, by itself do all the work, but the conceptual nature of perception can't either.

At this point we should examine in detail the latter thesis, but I think that we can see what is problematic in the conjunction of IT, ITP and CP without entering McDowell's theory of the conceptual content of perception. For perceptual thoughts, unlike possibly other kinds of thoughts, seem to be contingent, and it is the point of saying that they *can* be true or false, *thinkables*. We have to remind here our discussion of the version (a) of ITF, where the identification is of thoughts with possible or actual facts, in contrast to the (b) version, where thoughts are identified with essential facts, for the latter version commits us to saying that the relation between thoughts and facts obtains *necessarily*. McDowell's talk of thinkables commits him to the bipolar conception of facts and propositions ITFa. But our previous discussion of this view showed that it met serious problems with the notion of fact.

Could McDowell claim that perception involves a relationship with what are *essentially* facts, in the style of (ITFb)? His talk of perception as "openness" to facts, and his insistence that perception is a real source of knowledge, might lead us to think this. If so we can derive, trading upon the ambiguity between (ITFa) and (ITFb), an argument *against* his identity theory of perception (Dokic 1998). The argument is this:

- (a) If perception has a content, then it is a "bipolar" content
(susceptible of being true or false in the sense of ITF a)
- (b) But perceived facts are not bipolar (they present us with what are
essentially facts, hence necessarily true thoughts, in the style of ITF b)
- (c) Hence the content of perception cannot be identical to a fact

If McDowell rejects (a), he will be forced to withdraw his claim that the contents of thought are thinkables. If, on the other hand, he rejects (b), it will be difficult to understand how thinkables which simply *may* be true can involve a relation to something existent and necessarily external to us, for in the case where they might be false they obviously relate us to nothing in the world (or they would relate us to negative facts: not a very illuminating contact with reality). Either way, the identity theory of truth appears damaging for a sound defence of the identity theory of perception. [\[19\]](#) And if ITP is conceived as an argument for, and not simply as a special case of, IT, the falsity of ITP threatens IT itself.

It may be objected that this argumentation against the coherence of IT, ITP and CP rests upon an implicit reading of the notion of the truth of perceptual thoughts in an external or transcendental realist sense that McDowell intends precisely to reject when he presents his identity theory of truth as a “truism”, and that I have been abused by the “high flown language” in which it is dressed up. The identity theory of perception, combined with the identity theory of truth would thus just amount to this, in a deflationary mood: a (perceptual) thought that p is true (or is a fact) if and only if p . But then what remains of the insistence that in perception we are “open” to “facts” which are “manifest”? In particular how can perception be a source of knowledge of the external world which impinges on our senses if its deliverance is conceptual? This is of course a major worry in *Mind and World*, which McDowell faces under the guise of the “oscillation” or “seesaw” between the coherentist version of the justification of our empirical beliefs that Davidson proposes and the Myth of the Given. His solution, if I understand him well, consists in saying that experience gives us at the same time a picture of what is outside and a picture which is conceptual in nature. From within the realm of sense, or the space of reasons, we feel the *passivity* or, in Kantian terms, the receptivity, of experience as of an outer world. Then our passive impressions become experiences of an outer world by being taken as such (*that things are thus and so*) by the active faculty of understanding (Kant’s spontaneity), i.e by being integrated within a conceptual repertoire, and revised, within our conceptual sphere (1994: 29-40). But here, like several commentators^[20], we may suspect that we have come back to a coherence version of the justification, where experience, as conceptual, coheres with our beliefs. The only way for McDowell to escape this conclusion would be to revert to a correspondence version of truth, where experiences are what *make true* our perceptual thoughts. But this would be inconsistent with an identity theory of perception as well as of thought in general.^[21] A last alternative is left: the combination of a coherence theory of justification with an identity theory of thought. The conceptual contents of experiences would then be coherent *and* identical to facts. But this is strongly reminiscent of an identity theory of thought (IIT), and of Hegelian or a Bradleyan idealism.^[22] If what I said in the previous section about the permanent temptation of an identity theory of thought to nudge the world into the mind is correct, it is not perfectly clear that McDowell has given us any means to resist this temptation. Yet he says:

“Conceptual contents that are passively received in experience... are about the world, as it appears or makes itself manifest to the experiencing subject, or at least seems to do so. that ought not to activate the phobia of idealism.”(1994: 39)

But if we can get the benefit of externality only from within a conceptual sphere of the “space of reasons” which is “unbounded” and has no “outside”, at least the fear of idealism is still there.

So even if McDowell does not hold an identity theory of truth for thought in general, he holds such a theory for perceptual thought. But then either this amounts to a modest or minimalist kind of identity theory, which it is difficult to distinguish from a minimalist theory of truth *simpliciter*, or he intends a more robust conception, but then the latter encounters the same problems as ITF and ITT above. Either way, this poses problems for his view of perception: for a minimalist IT can hardly make sense of perception putting us into contact with an independent reality, and a substantive theory may undermine the conceptual character of perception. [\[23\]](#)

5. Hornsby’s “identity theory”

Jennifer Hornsby (1997) has recently defended a view of truth which she calls “the identity theory”. She takes her departure from McDowell’s claim quoted above about the identity of true thinkables and facts, which she takes, like him, as a truism, which should not encourage any substantial theory :

“For sure, the identity theorist introduces the word ‘facts’, but that is only in order to find a way of saying that the facts are the same as what is true, and thus be quite explicit about her opposition to a correspondence theory of truth.”(1997: 20)

We might wonder whether the identity theory here proposed as distinctive differs from the deflationary theory. But Hornsby also denies that what she calls the identity theory amounts to a deflationary view (p.21), and she says that it is “not vacuous” (p.3) and has “definite commitments”, although it “embodies nothing metaphysically contentious” (p.9). What are these commitments which purport to make this a genuine alternative option?

As far as I can see, there are mainly two such commitments. [\[24\]](#) The first one is that of taking truth bearers as thinkables. This implies that there are contents of thoughts and meaning, and that HIT is not reduced to a disquotationalist view where truth bearers are sentences. Thinkables, according to Hornsby, are essentially the contents of thoughts that are interpreted under the constraints of an interpretation theory in Davidson’s style (p.12). Understanding what an other speakers thinks is to interpret the content of thought as a thought which is possibly true, and about the world. We can, she says, also extend the notion to the yet unthought thinkables of humanity as ideal possible facts encountered in an ideal situation. The second commitment is related to the first: if thinkables are meanings, then there is such a thing as a what a theory of meaning is a theory of. This is the point where Hornsby’s

identity theory diverges from the deflationary theory, for a deflationary conception of truth should lead us to a deflationary theory of meaning (along the lines indicated by McDowell quote from his (1981) above, for instance). On the contrary, an interpretative account in Davidson's style makes "these deflationary impulses go away"(p.19).

Each of these commitments seems to me to raise problems. First, if true thinkables are facts this raises exactly the same kind of difficulties as those that we encountered with McDowell's use of this notion: to repeat, if they are bipolar and identifiable with either possible or actual facts, how can truth bearers be identical with the truth makers, and if they are essentially facts, shall we say that false thinkables are negative facts? Hornsby does not answer these questions, presumably because she does intend to defend an ITF in the sense of § 2 above. If she is not, it is difficult to see whether the use of the notion of fact is not purely deflationary. The difference between an identity theory of truth couched in terms of the deflationary notion of fact and a deflationary theory of truth would thus amount to the difference between the latter and an interpretative account of thought in Davison's style. The claim, however, that thinkables are to be uncovered by an interpretation theory is not presented by Hornsby as purporting "to *explain* either truth and meaning" (p17), but as putting us in position "to elicit features of one's conception of truth" (*ibid*). So this conception would help us both to resist a deflationary or modest attitude towards truth and modesty towards meaning.

This is an interesting proposal, which I do not want to discuss here, but it is unclear that the attempt to find a space between substantive conceptions of meaning and truth and deflationary conceptions of them has anything to do with an identity theory of truth, at least in the various senses which we have examined here. What comes closer, in Hornsby's presentation of the "identity theory" to the views that I have discussed is the deflationary conception of facts. But it is not, have we have seen, an identity theory of truth in any of the other, substantial senses, since it is just a version of the deflationary conception. Hornsby's, however, insists that her "identity" theory "cannot be vacuous because it takes a stand on what the bearers of truth are, calling them thinkables". But she also agrees that this stand is independent from both a substantive conception of independent facts and from an uninteresting trivial conception of them. The matter may be purely terminological, but it has nothing to do with the notion of identity which gives its name to the proposed theory. It is a claim about the nature of truth bearers, whereas all the different versions of the identity theory that we have distinguished are also claims about truth makers in their relation to truth bearers. If one wants to insist, like Hornsby, that we have to avoid the idea that there are entities which stand in the truth making relation with to truth bearers by

saying that truth is predicated of thinkables, this is a quite legitimate claim, but the notion of identity does not by itself fulfill this job.

6. Conclusion: through the looking glass

I am thus sceptical about the pretensions of an identity theory of truth to offer a genuine third way between the usual sorts of substantial and minimalist theories of truth. In all its possible versions, it oscillates between a truism about the identity of propositions and facts which makes it close to the deflationary view and substantial conceptions which imply commitments which are far from coherent. This does not mean that the deflationary view is all that we can get, and that the urge felt by identity theorists to expand the deflationary view is not legitimate.^[25] But it is misplaced. Why, nevertheless, is an identity theory attractive? We can see better why. On the one hand, truth is, as Blackburn says, “transparent”: “It is as though you can always look through “it is true that” to identify the content judged, inquired after, and so on as if the reference to truth were not there” (Blackburn 1984: 227). We can stick to this feature, and to a deflationary view. But we may also try to explain it, and feel the “obstinate instinct” of identification of truth with reality that Bradley talks about. It is as if we could see the world through our truths, as if we could have the benefit of the external world from within our own sphere, with a sort of cosmic familiarity. But, to paraphrase McDowell (1981), the right course is to set our faces against the idea of cosmic familiarity.

References

- Armstrong, D. 1997 *A World of States of Affairs*, Cambridge, Cambridge University Press
- Baldwin, T. 1990 *G.E. Moore*, Routledge, London
- 1991 “The Identity Theory of Truth”, *Mind*, 100, 35-52
- Barwise, J. 1989 *The situation in Logic*, Stanford, CSLI
- Blackburn, S 1985, *Spreading the Word*, Oxford, Oxford University Press
- and Simmons eds, 1999, *Truth*, Oxford, Oxford University Press
- Bradley, F.H. 1893 *Appearance and Reality*, Oxford, Oxford University Press
- Candlish, S. 1989 "The Truth about F.H. Bradley", *Mind*, 331-48
- 1996 “Truth, identity theory of”, *Stanford Encyclopedia of Philosophy*, [http:// plato.stanford.edu/ entries/truth-identity](http://plato.stanford.edu/entries/truth-identity)
- 1999 “Identifying the Identity theory of truth” *Proceedings of the Aristotelian Society*, XCIX, 233-240
- 1999 a “A Prolegomenon to the Identity Theory”, *Philosophy*, 74, 288, 199-220
- Cartwright, R. 1987 “A Neglected Theory of Truth”, *Philosophical Essays*, Cambridge Mass, MIT Press
- Dodd, J. and Hornsby, J. 1992 'The Identity Theory of Truth: a Reply to Baldwin', *Mind*, 101, 319-322
- Dodd, J. 1995 " McDowell and Identity Theories of truth", *Analysis*, 55, 162-165

- 1999 "Hornsby on the Identity Theory of Truth", *Proceedings of the Aristotelian Society*, XCIX, 225-232
- Dokic, J. 1998 "The Ontology of Perception: Bipolarity and Content", *Erkenntnis*, to appear
- Fine, K. 1982, "First-Order Modal Theories III- Facts", *Synthese*, vol. 53, 1, 43-122
- Frege, G. 1967 "The Thought", in Strawson, P. ed. *Philosophical Logic*, Oxford, Oxford University Press
- 1979 *Posthumous Writings*, Blackwell, Oxford
- 1903 *Grundgesetze der Arithmetik*, vol. 2, Jena, repr. Olms
- Friedman M. 1996 "Exorcising the Philosophical Tradition: Comments on John McDowell's *Mind and world*", *The Philosophical Review*, 4, 427-467
- Hornsby, J. 1997 "Truth: the Identity Theory", *Proceedings of the Aristotelian Society*, XCVII, 1-24
- 1999 "The Facts in Question: a Reply to Dodd and Candlish", *Proceedings of the Aristotelian Society*, XCIX, 241-245
- Horwich, P. 1990 *Truth*, Oxford, Blackwell second ed. 1998
- McDowell, J. 1981 "Antirealism and the Epistemology of Understanding", in Bouveresse, J. & Parret, eds. H. *Meaning and Understanding*, De Gruyter, Berlin, 1982 and in McDowell 1999
- 1987 "In defense of modesty", in B. Taylor, ed *Michael Dummett, Contributions to Philosophy*, La Haye, Nijhoff, 59-79 and in McDowell 1999
- 1994 *Mind and World*, Harvard, Harvard University Press (*quoted after the first edition; second edition 1998*)
- 1998 *Mind, Value and Reality*, Harvard, Harvard University Press
- 1999 *Meaning, Knowledge and Reality*, Harvard, Harvard University Press
- Sedgwick, S. 1997 "Mc Dowell's Hegelianism", *The European journal of Philosophy*, 21-38
- Wittgenstein, 1913 *Notes on logic*, in *Notebooks*, tr. Anscombe, Blackwell, Oxford 1961
- Wright, C. 1992 *Truth and Objectivity*, Harvard, Harvard University Press
- 1996 "Human Nature?" review article of *Mind and world*, *The European Journal of Philosophy*, 4.2, 235-253
- 1996 a "Responses to commentators" in *Philosophy and Phenomenological Research*, LVI, 4, 911-941.
- 1998 "Mc Dowell's oscillation", *Philosophy and Phenomenological Research*, LVIII, 2, 395-403.

-
- [1] See for instance Horwich 1990, Wright 1992, and the anthology of papers by Blackburn and Simmons (1999). For my purposes here, I do not distinguish between various kinds of minimalist theories. I just take minimalism or deflationism as the view that the disquotational schema exhausts what we can say about truth, which cannot be an explanatory property.
- [2] I say in part, because Hornsby does not intend to propose a substantive theory in terms of facts. See below, §5
- [3] see e.g. Blackburn 1985, 255 ff.
- [4] Frege 1967: 20, Frege 1979: 129, Frege 1903: 245 (see Dodd & Hornsby 1992: 320)
- [5] I take these phrases from Candlish 1999: 206-211
- [6] Frege 1980: 63
- [7] Baldwin 1991: 47
- [8] Wittgenstein 1913: 196. I am here indebted to Dokic (1998).
- [9] See e.g. Fine 1982, Barwise 1989
- [10] One of the familiar reasons why the notion of negative fact is worrying is that, to put it in terms of an identity theory, a single truth would also be identical with an indefinite number of such negative facts.
- [11] Dokic 1998 remarks that this is a version of “disjunctivism” about thought (in parallel with the disjunctive theory of perception) : either a thought is a fact, or not; whereas ITFb (i)-(iii) is a version of conjunctivism about thought : two kinds of thought, true ones and false ones, two kinds of facts.
- [12] As Baldwin notes about Moore: “The resulting metaphysical system can seem almost idealist: the world is, quite literally, a world of meanings.”(Baldwin 1990: 42-43)
- [13] Baldwin 1991: 36
- [14] Bradley 1893: 319-320
- [15] See Horwich 1990 111-112, Wright 1992, 25-26
- [16] *Principles of Logic*, 2nd ed. Oxford 1922, p.49-50 quoted after P.Hylton, *Russell, Idealism and the Emergence of analytic Philosophy*, Oxford, Clarendon Press 1990.
- [17] Dodd 1995 and Candlish 1999 at least attribute an IT to her, although Hornsby (1997, p.2, note 2), who takes her departure from McDowell to formulate her own version of IT, says that she is unsure whether McDowell intends to be an identity theorist
- [18] It has to be added “when we are not misled”, i.e. when perception is veridical. This may seem to beg many questions, but I shall not enter into this.
- [19] Dokic (*ibid*) further remarks that we could reverse the argument (a)-(c) above by arguing that since the content of perception is the same as that of judgement, hence conceptual, and conceptual content is bipolar, the perceived fact is also bipolar. It is too thinkable. But it remains to be seen how this can be consistent with perception being a source of knowledge, or true contents, which is just the point examined in the following paragraph.
- [20] Friedman 1996: 443-444; Wright 1996a: 240-42 : “A reader might well wonder how anything essentially at odds with Coherentism has been proposed – since at that may seem to have been effected in an enlargement of the terms of the coherence relation. Before, we thought of coherence as essentially a relation on beliefs. Now for basic empirical belief, we impose an additional requirement : coherence with experience, with the latter conceived as content-bearing after McDowell”. See also Wright 1998: 397.
- [21] At p.179 of 1994, McDowell considers explicitly the objection that we have met above about the attribution of an IT to Frege, that it would fudge the realm of sense into the realm of reference (see also Dodd 1995). But he rejects the objection by saying that his taking thinkables to be the bearers of truth implies that Frege’s notion of sense is the notion in the context of which “we should reflect about the relation of thought to reality, in order to immunise ourselves against the familiar philosophical anxieties” (p.180). But whether we agree or not that for Frege having a sense implies having a reference, the relation of thought to reality cannot for him be secured only at the level of sense.
- [22] Indeed McDowell himself presents his book as a prolegomenon to Hegel ‘s *Phenomenology of spirit*. Many commentators would not mind at this conclusion. See for instance Sedgwick 1997, who argues that McDowell is consistently Hegelian.
- [23] I reach conclusions similar to those of Dodds 1995, but with respect to perception more specifically.
- [24] There is a third commitment : that “the conception of truth which the identity theory brings with it allows truth to be a *sui generis* norm, in play where there are rational beings who may go right or wrong in their thought or speech.”(p.23). But I shall leave it aside here.
- [25] A conception of “minimalism” like that defended by Wright (1992) seems to me legitimate.