For an issue on Concepts, edited by E. Lalumera, of *Dialectica* 64 (2010) 35-56

Relation-Based Thought and Its Philosophical Significance[1]

Christopher Peacocke

For a kind of thought about a subject-matter to be *relation-based* is for thought of that kind to be available to a thinker only because of certain relations in which the thinker stands to the subject-matter in question. There are ways of thinking of a particular object that are available to a thinker only because of a certain relation in which the thinker stands to the object. There are ways of thinking of a particular property that are available to a thinker only because of a certain relation in which the thinker stands to some instances of the property; and so forth. Some examples of relation-based thought are very familiar, and they continue to be a proper target of philosophical investigation. These familiar cases include demonstrative ways of thinking of objects and events given in perception. These ways of thinking are given linguistic expression, in context, by the utterance of such phrases as 'this tree', 'that fire'. Another familiar category of relation-based thoughts consists of those containing concepts made available by the thinker's perceptually-based recognitional capacities for objects and properties. It is well-known that wildly false, inadequate, and incomplete theories of something are consistent with thinking of it in a perceptual-demonstrative or a recognitionally-based way. My principal

[1] This short paper was written for the Workshop on Objectivity and Disagreement held at the New York Institute of Philosophy, at New York University, on 1 December 2007. With the aim of promoting discussion, it was written as the philosophical analogue of a 'position paper' for a political meeting. In the issues of almost every paragraph, there are major alternatives to those propounded here, and many more important turns to the arguments than those described here. At the New York meeting, I received illuminating comments from many participants, and from my commentator Paul Horwich. Rather than replying and producing a very different kind of paper, in the present version I have retained the primary aim of promoting discussion, now in the wider philosophical community.

claim in this paper is that the phenomenon of relation-based thought and its significance vastly outrun these perceptual cases.

My aims are to present a case that relation-based thought has this far more extensive character; to offer an explanation of how these more extensive cases are possible; and to trace out the consequences of this explanation for the scope and limits of intelligible disagreements between thinkers. I will be making some specific comments on spatio-temporal thought, thought about conscious states, thought about meaning and content, and thought involving logical constants. Although I obviously cannot be talking about all the distinctive kinds of subject-matter about which there may be substantial disagreement – for that seems to cover pretty much all the subject-matters there are - I believe that some of the explanatory models I offer can be generalized to other areas.

I propose the following theses about relation-based thought. You will be pleased to learn that, unlike Martin Luther (who had ninety-five theses), I will be posting only four:

> (I) The phenomenon of relation-based thought is extensive, running far beyond the familiar perceptual examples. It is particularly salient amongst concepts of events and states in the conscious realm.
>
> (II) In a range of cases, a relation-based way of thinking of a particular property involves tacit knowledge that its instances stand in a certain relation to some distinguished instances picked out by their special relation to the thinker. When understanding takes this relational form, no particular (non-trivial) relations of evidence or consequence have to be grasped for full understanding of the notion in question to be present.
>
> (III) Ordinary thought about particular meanings and intentional contents - thought about these meanings and contents as the meanings and intentional contents they are - is itself an instance of the phenomenon of relation-based thought.

(IV) The generalization to all cases of the idea that grasp of a concept consists in tacit knowledge of its fundamental reference rule can explain both constitutive and psychological facts involving the rationality of judgements involving the concept. It can explain the minimal objectivity that judgement enjoys; it can provide some substantial constraints on the intelligibility of revisions of the principles we accept; and it can explain some otherwise puzzling phenomena involving the logical concepts.

If the first three theses are true, then there will be a wide range of concepts, including those of specific meanings and intentional contents, whose nature is not given by any kind of grasp of their role in theories of particular kinds. If the fourth thesis is true, a proper appreciation of this fact, far from precluding us from giving a positive account of understanding, suggests a model of understanding that accounts for objectivity, and accounts for the nature of rational acceptance and the possibility of rational revision of the principles we accept.

I

First, then, for Thesis (I), that relation-based thought is far more extensive than the perceptual examples. Concepts of all of the following mental event-types and states each seem to me to display some variety of relation-based thought:

pain and other bodily sensations;
visual, auditory and other sensory experiences;
occurrent conscious emotions;
mental actions.

The occurrence of a pain-event to a subject can make it rational for her to judge ‘I’m in pain’. The occurrence of a visual experience of a desk in front of her can make it rational for her to judge ‘I have an experience as of a desk in front of me’; and so forth. It is not merely that these conscious experiences and other mental events make these first-person

present-tense judgements rational. It is, further, constitutive of adequate possession of the concepts *pain*, *visual experience* and the rest that the thinker be willing rationally to apply the concepts in response to the instances she experiences. What makes these concepts the concepts they are is in part this rational sensitivity of judgements containing them to certain occurrences of the very events of the sort they pick out.

It follows that we have here (part of) an account of how the thinker's concept of these conscious states latches on to them, rather than to some other mental states and events. The thinker's concept latches on to the right mental states and events without relying on the thinker having any conception, either explicit or tacit, of the role of events of these mental types in her own, or in others', psychological economies. Many thinkers have demonstrably false beliefs about the role of their own pains and their own visual experiences without this casting any doubt on the fact that it is indeed their pains and visual experiences that they are thinking about. Some subjects think that they withdraw their hands from hot radiators because of the pain they experience on touching the radiator. In fact the withdrawal is a reflex. In some cases the withdrawal of the hand is initiated even before the subject experiences any pain. Again, many subjects think that their bodily actions are guided by their visual experiences. In fact we know that much such action is caused by the more rapid unconscious representations of the environment served up to the brain by the older, dorsal route.[2] But subjects who have this false belief are nevertheless having a false belief about visual experience. What makes it so is the rational responsiveness of their concept of visual experience to the occurrence of their own visual experiences. In short, these mental concepts are individuated in part by the relations of one who possesses them to certain instances of the concepts themselves. So these are all cases of relation-based thought. This is a species of relation-based thought that, in my judgement, legitimates the description of these mental concepts as ones of which the thinker knows from her own case what it is for them to apply to an event or state.

[2] See A. Milner and M. Goodale, *The Visual Brain in Action* (Oxford: Oxford University Press, 1995) and M. Goodale and A. Milner, *Sight Unseen* (Oxford: Oxford University Press, 2004).

That description may well set several alarm bells ringing. They will ring in the minds of those who believe that the Myth of the Given has been refuted. They will also ring in the minds of those who believe that there is something right in the Private Language Argument that blocks any such conception of the self-ascription of conscious states. The theses underlying such alarm merit essays of their own. Here I will briefly indicate what seems to me to be a position on these issues that is consistent with what I have endorsed so far.[3]

A transition by a thinker from

the occurrence to her of a mental event that does not involve any content about the world

to

a judgement by her about the nonmental, objective world

cannot by itself possibly be a rational transition. Any such transition would be a nonrational leap in the dark. With that much of the critical discussion of the Myth of the Given I am in agreement. Nothing I have said contradicts it. When a thinker self-ascribes pain in rational response to pain itself, the judgement she makes is not about the world beyond the pain itself. Its truth requires no more than the occurrence of the state that makes the judgement rational. This is not a leap in the dark, but an entitled, paradigmatically reasonable step. If it is thought that it must be a nonrational step because it is not a rational response to a conceptualized state of the subject, what these simple considerations seem to show is that the demand that the rationalizing state be conceptualized is too strong.

The problematic of the Private Language Argument has been even more extensively discussed. For some, such as McDowell, it is connected with the Myth of the

[3] On 'The Myth of the Given', see W. Sellars, *Empiricism and the Philosophy of Mind* (Cambridge, Mass.: Harvard University Press, 1997) sections 1-7, pp.13-25; R. Brandom, *Study Guide* in the preceding volume, pp.120-131; J. McDowell, *Mind and World* (Cambridge, Mass.: Harvard University Press), Lecture I.

Given.[4] Here I can say only that under what I have endorsed so far, there is a clear distinction between someone who is really following the correct rule for applying a word which means *pain*, and someone who merely thinks she is. It is the distinction between someone whose use of the word is keyed to her really being in pain, and someone for whom that is not the case. This description of the distinction involves nothing about criteria, and does not, in the first instance at least, immediately invoke anything about knowability.

This feature of concepts of conscious states that ties their individuation to a relational property of self-ascriptions cannot be an exhaustive account of the nature of these concepts. It says nothing about what it is for a third-person ascription of the concept to be true. The rational sensitivity of first-person ascriptions of a given concept *C* of a conscious state to instances in the thinker of *C* is not even a feature which is unique to the concept *C*. The feature will equally be present, in the case of the concept of visual experience, for example, in a concept that is stipulatively restricted to the thinker's own visual experiences. The feature will also be present in a concept of which it is indeterminate whether it applies to the visual experiences of subjects other than the thinker. Both of these concepts are distinct from that of our actual concept of visual experience, of which it is determinate that it can apply in the third-person case. The same applies to all of our other actual concepts of conscious states and events.

So our next three questions should be:

(a) What is the correct account of our understanding of the third-person case?
(b) How is it integrated with this relation-based account of the first-person case?
(c) Does a plausible account of the third-person case also support the description of it as relation-based thought?

These questions are special cases of a form of question that arises equally for many other cases in which we have a relation-based way of thinking that individuates a concept in part, but not wholly, in terms of the relation of some thoughts containing it to certain local applications. For example, recognitional observational concepts of shape are

[4] *Mind and World* pp.18-23.

plausibly individuated in part by their relations to perceptual experiences that represent things as being of a certain shape. But many shape concepts can be truly predicated of things not currently, and perhaps not ever, perceived (and perhaps not even perceivable by us, in the case of the extremely small and the extremely large). So here too we must explain our understanding of the non-local case, and address the corresponding version of the questions (a) – (c).

II

Understanding connects the non-local instances of a concept with the local ones by means of tacit knowledge of a condition involving the relation of identity. For a non-local object or event to have the property picked out by a concept is for it to have the same property (of some general kind) as the local cases when they fall under the concept. One who grasps the concept tacitly knows this identity-involving condition for a non-local object to fall under the concept.

This abstract general form is realized in various different ways specific to various different subject-matters. Here are some examples:

> For an arbitrary object to fall under the observational concept *round* is for it to be of the same shape as things are represented as being in perceptual experiences of objects as round, experiences which make it rational to judge 'That's round' in the local case.

> For an arbitrary event to fall under the concept *pain* is for there to exist some subject of consciousness who experiences that event and for the event to have the same subjective property for that subject as is possessed by events that make rational the application of the concept *pain* when you judge 'I'm in pain'.

There are strict preconditions on this sort of thing - explanations of understanding in terms of grasp of identity - making sense (as Wittgenstein would have insisted). In the spatial example of the observational concept *round*, it essential that we have a theory of

perceptual content under which it is genuinely a region of space itself in which something perceived as round is presented. If this condition were not met, then it would not make sense, or at least could not ever be true, that some arbitrary region of space is of the same shape as a region in which something perceived as round is presented. In the case of concepts of conscious states, there are equally preconditions.

There are equally preconditions in the case of concepts of conscious psychological states and events. We need to have available a conception of multiple subjects of experience, of which one is oneself only one, for the explanation of understanding of concepts of conscious states to make sense, or to give the correct extension for the concept. An account under which only your experiences are can be counted as subjectively the same as any experience of yours would prevent the identity-involving explanation from determining the correct extension of the concept. My view is that these essential preconditions can be shown to be fulfilled, though it is certainly a substantial philosophical task to do so. A paper showing how these important preconditions are met would have different goals from this one: here I am just acknowledging their existence.[5]

The very form of these explanations of understanding in terms of an identity relation gives them two attractions. One is that the form of the explanation ensures that the meaning of the local predications of the concept in question and the non-local predications are univocal. They are saying the same thing about the local and the non-local case. The univocality is ensured by the requirement that it be the same property in the non-local case as in the local case that is required for the truth of the respective predications.

Other substantive theories of understanding have had great difficulty in meeting the requirement of univocality. Neo-Wittgensteinian theories of psychological concepts commonly treat third-person predications in terms of the fulfillment of certain criteria, and treat the first-person predications utterly differently, usually as merely expressive. I would argue that these neo-Wittgensteinian theories have never been able to demonstrate the univocality of the predications of concepts of conscious states under the treatment of

[5] I attempt to argue that these preconditions are fulfilled, and try to explain why, in *Truly Understood* (Oxford: Oxford University Press, 2008), Chs. 1 and 5.

understanding which they offer. The same issues arise, of course, for the relation between first-person present tense predications of conscious states and first-person predications in other tenses.

The other attraction of the general form of explanations of understanding in terms of tacit knowledge involving an identity is that they immediately conform to the Fregean conception of a sense as fixed by the fundamental condition for something to be the reference of the sense. The identity-involving understanding condition states what is required, as a matter of the identity of the concept, for an arbitrary object or event to fall under the concept *pain* or *round*. By fixing this condition, we fix the sense under the classical Fregean conception of the *Grundgesetze*.

The account of understanding of the non-local case in terms of tacit knowledge of an identity connecting it with the local case is not an account in terms of conceptual role. It is not an evidential account. It is not an account in terms of reasons. It is not an account in terms of consequences. In my judgement, neither evidence, nor specific reasons, nor specific consequences should be mentioned in an account of understanding what it is for tiny object to be round, or for some other person or organism to be in pain. If you have tacit knowledge of the identity condition required for understanding these non-local predications, then you know what it is for the predication to be true. It may remain a completely open question for you what might be evidence, here and now, that that condition is fulfilled, or what the consequences, here and now, are of that condition's being fulfilled. There is no a priori connection between any particular kind of evidence - nor any particular kind of consequence, noncircularly specified - and the tiny object being round, or the other organism's being in pain. What would be such evidence has to be worked out empirically, in combination with the tacit knowledge involved in understanding. It may take ingenuity and imaginative reasoning to do so. Understanding is explanatorily prior to evidence and consequence in these cases.

Correlatively, widely divergent assessments by two thinkers of what would be evidence for, or the consequences of, these non-local conditions holding does not need to cast doubt on their possession of a common understanding. If two thinkers have the relevant tacit knowledge of the same identity-condition in their understanding of what it is for something to be true in the non-local case, they will be disagreeing about the

evidence for, or consequences of, the same propositions or Thoughts containing the same concepts.

The general conclusion of this section is then that there is a second sort of relation-based thought. Unlike the first sort that involves some form of relation of consciousness to instances of the property in question, this second sort involves rather a piece of tacit knowledge involving an identity relation, a piece of knowledge about the correctness conditions of the application of a particular concept in non-local cases; and this second sort correspondingly allows for a much wider scope of intelligible disagreement about the subject-matter in question.

III

The apparatus and approach I have introduced so far can be applied to those concepts of particular concepts and particular meanings that present those concepts and meanings as the particular concepts and meanings they in fact are. When you think of a concept as 'the concept *man*', and fully grasp that concept of a concept, you know which concept you are thinking about. This concept of the concept *man* is made available by your possessing the concept *man* itself. If you are capable of thinking about concepts of concepts, and possess the first-level concept *man*, then you are in a position to possess the second-level concept 'the concept *man*'.

The converse seems also to hold. There seems to be no such thing as fully possessing the relevant second-level concept 'the concept *man*' without possessing the first-level concept to which it refers.

We can speak of this special second-level concept as the canonical concept of the concept *man*. In general, we can speak of the canonical concept of the concept F of one level lower, and we can use the notation 'Can(F)' for the canonical concept of the concept F. It seems that the special relation between a thinker and the concept *man* that makes available to her the canonical concept of that concept is simply the relation of possessing that first-level concept *man*. But how does it do so? What is the explanation of this phenomenon, and what is its significance?

There would be a very straightforward explanation of the phenomenon under a pleonastic theory of concepts, in Stephen Schiffer's sense.[6] According to the pleonastic treatment of concepts, a sentence such as

(1) Smith falls under the concept *man*

is simply a stylistic variant of

(2) Smith is a man.

The pleonastic conception holds that it is wrong to regard the former sentence as importing or invoking a substantial ontology of concepts. Schiffer's theory would certainly explain the data I just cited. It is a virtue of the pleonastic treatment that it immediately implies that there is no grasping the canonical concept of the concept *man* without possessing the concept *man*. If grasping the canonical concept of the concept *man* involves appreciating that (1) is no more than a stylistic variant of (2), that datum immediately follows. The converse datum is also explained. If a thinker possesses the concept *man*, and has the understanding of the operation of phrases of the form 'the concept of ….' provided by the pleonastic account, the thinker will have all that is required to possess the canonical concept of the concept.

I want, however, to offer an explanation of the phenomenon that is also consistent with a theory of concepts that allows the explanation of our understanding of sentences of natural language to be explained by our grasp of the concepts expressed by the component words of the sentences, together with the significance of their mode of combination. Schiffer's account of concepts would not allow concepts and conceptual combination to be explanatory of understanding in that way. Schiffer and I diverge over the acceptability of this consequence. He regards it with equanimity, and does not (or did not) think that an explanation of linguistic understanding needs to be compositional in

[6] On pleonastic conceptions of properties, facts and propositions, see S. Schiffer, *Remnants of Meaning* (Cambridge, Mass.: MIT Press, 1987), 51; and his *The Things We Mean* (Oxford: Oxford University Press, 2003), esp. Ch. 2, 'Pleonastic Properties'.

that way. I myself do hold that an explanation of linguistic understanding needs to be compositional in that way. I will not reiterate those arguments here. I want instead to show how the apparent data about canonical concepts of concepts can be explained equally under an approach that is consistent with, and indeed implies, a compositional theory of meaning and understanding.[7]

The generalization of the approach I have been adopting is one under which a concept is individuated by the fundamental rule for something to be its reference, that is, the rule that states what makes some entity of the appropriate category its reference. As in the previous cases we considered, here also to possess the concept is to have tacit knowledge of the fundamental reference rule for the concept. Applying this approach, I suggest this fundamental reference rule for the canonical concept of the concept C:

(FRR Can(C)): For an arbitrary concept φ to fall under Can(C) is for φ to have the fundamental reference rule that an arbitrary object x falls under it iff R(x), where this last biconditional is in fact the fundamental reference rule for the concept C.

It follows from this that any thinker who has tacit knowledge of (FRR Can(C)) must also possess the concept C itself of which Can(C) is the canonical concept. To have tacit knowledge of the displayed rule requires the subject to represent the condition of an arbitrary object falling under a given concept iff it is R, which last is itself the fundamental reference rule for C. But under this approach, that is for the thinker to possess the concept C. So the approach explains why anyone who has the canonical concept of a concept C must have the concept C itself.

---

[7] For an exchange between Stephen Schiffer and me on these issues more years ago than I care to count, see his *Remnants of Meaning* Ch.7, "Compositional Semantics and Language Understanding"; my paper "Explanation in Computational Psychology: Language, Perception and Level 1.5" *Mind and Language* **1** (1986) 101-123; Schiffer's article "Peacocke on Explanation in Psychology" *Mind and Language* 1 (1986) 362-371; and my Reply in the same journal issue 395-7. For a discussion of Schiffer's position in Chapter 2 of *The Things We Mean*, see section 26.6 of my 'Concepts and Possession Conditions', in *The Oxford Handbook of Philosophy of Mind* (Oxford: Oxford University Press, 2009), ed. B. McLaughlin and S. Walter.

Conversely, if a thinker possesses C, she must be able to represent its fundamental reference rule R. So if she has the apparatus to think of concepts of concepts, she has all she needs to formulate and grasp (FRR Can(C)); which is to say that she is in a position to possess the canonical concept of C as well.

Under this approach, the biconditional

(3) Fa iff a falls under the concept F

is something that we can derive as a consequence of the fundamental reference rule (FRR Can(C)) for the canonical concept of F. The derivation from left to right of the biconditional (3) runs:

Fa
So, a meets the condition R, where R is the fundamental reference rule for F;
So, a falls under the concept whose fundamental reference rule is R;
So, by (FRR Can(C)), a falls under Can(F);
which is to say, a falls under the concept F.

The derivation from right to left simply reverses this derivation (with corresponding adjustments of the rationales for each step). In short, what the pleonastic theory treats as a matter of stylistic rewriting is here regarded as a substantive equivalence, something in need of a (simple) proof.

The status of (3) under the present treatment is in some respects analogous to the status of disquotational T-sentences of the form

'A' is true iff A

in a theory of truth for a language. Like instances of (3), these T-sentences are obvious (if we prescind from the paradoxical and vacuous cases). All the same, the obvious T-sentences need to be given a derivation from the reference-conditions of the expressions in A if we do not hold a minimalist theory of truth, and if we want to defend a

compositional and truth-conditional theory of understanding. Similarly, the instances of (3) also need to be given derivations from an account of what it is to be the canonical concept of a concept, if we do not hold a pleonastic theory of concepts.

Part of the philosophical interest of this account of the canonical way of thinking of a concept, or equally of a meaning, is that it helps to explain how a thinker can radically mischaracterize the nature of his own concepts and meanings, whilst still fully possessing those concepts and grasping those meanings. What underlies the ability to think of something as the concept F is a piece of tacit knowledge and, as with any other piece of tacit knowledge, a thinker may possess it whilst mischaracterizing it in his explicit, non-tacit theories and beliefs. Two linguists may disagree in their explicit theories about the syntax and semantics of some construction in a common natural language that they both fully understand. Similarly, two philosophers may disagree about the nature of a concept they both fully possess. The important point is that to think of a concept as the concept F, for example to think of a concept as the concept *and*, or the concept *round*, or the concept *perception*, is not to think of it as meeting the condition that in fact makes it the concept it is. That condition can be discovered only by extensive investigation and inference to the best explanation of a range of reason-involving phenomena of thought and judgement. Rather, to think of a concept as the concept *and*, or the concept *perception*, is to think of it in a way made available by the thinker's own possession of those very concepts themselves, whatever the thinker's conscious views, correct or incorrect, may be of the nature of those concepts. (FRR Can(C)) is meant to give the nature of this distinctive way of thinking of a concept. (FRR Can(C)) does not at all involve any explicit knowledge, on the part of the thinker, of what individuates the concept that he is thinking about in the canonical way that is its subject-matter.

Though it is no doubt all too clear that this paper is written from the standpoint of a defender of a classical truth-conditional theory of meaning and understanding, it is only fair to note that this most recent point is available equally to conceptual-role theories of meaning, in their many variants. If possessing a particular logical concept is grasping a particular conceptual role, that is apparently entirely consistent with a philosophical thinker mischaracterizing what that role is (just as he may mischaracterize his own practice of syntactic classification of sentences). For a particular role to be operative in a

thinker's ordinary judgements involving a particular concept, including a logical concept, is one thing. For a thinker to characterize that role correctly in his philosophical thought about his thought is quite another. Getting the description of the role right can take hard thought, and mistakes are possible, without that fact casting doubt on the thinker's grasp of the first-level logical or other concept in question.

IV

In this final section, I turn to the question: what more generally should our account of the role of understanding in rational judgement be? Can we generalize the model of understanding as a kind of tacit knowledge that has been present in the kinds of concepts we have considered so far? If we can so generalize the model, what is the role of understanding so conceived in rational judgement? Can that role explain the respect in which judgement has a certain minimal objectivity? How does it constrain intelligible disagreement? And what are the consequences of this account for the understanding of logical constants in particular?

Suppose a thinker judges that *p* in given circumstances, and that she makes the judgement rationally. The thinker will have certain reasons for her judgement. The judgement may be made as a result of inference from certain premises; but the thinker can have also have reasons without the judgement being made as a matter of inference. The reasons may for instance involve perception, memory, sensation or action-awareness. For the judgement to be a rational judgement, there must be some explanation of why these states and events, or any inference on which the judgement rests, provide reasons for the content *p* that is judged. This explanation must connect these states, events or premisses of the inference, and the form of the inferential or other transition, with the conditions for truth of the content *p*. They can be good reasons only if they are reasons for thinking this truth-condition is fulfilled. To put it in a slogan, what makes something evidence must be founded in the nature of what it is evidence for. Under any truth-conditional conception of intentional content, this task of explaining the status of reasons as good reasons never reduces to simply noting the consequences of the conceptual role

of the content *p* in question. That role, according to the truth-conditional conception, is not the way the content is individuated.

There are actually two tasks to be distinguished here. There is the philosophical task of explaining what is involved in the judgement's being rational. But since thinkers are broadly, though of course by no means universally, capable of being sensitive to the distinction between rational and non-rational judgements in given circumstances, there must also exist some psychological explanation of how thinkers themselves are able to draw this distinction, an open-ended distinction that applies to arbitrarily many contents and kinds of circumstance in which they may be judged. So there is also the task of saying how thinkers are capable of being sensitive to the distinction between the rational and the non-rational.

The resources to be drawn upon in answering these two questions are the two ideas of concepts as individuated by their fundamental reference rules, and of grasp of a concept as consisting in tacit knowledge of its particular fundamental reference rule.

To show that a mental state, for instance, really gives reason for judging that *p* in given circumstances involves showing that the thinker's being in the states, in the circumstances, makes it sufficiently likely that the truth-condition for *p*, holds, where this truth-condition is determined by the fundamental reference rules for its constituent concepts and their mode of combination. 'Sufficiently likely' is may be only one way of implementing this general conception (I am not wedded to particular implementations, only to the general conception itself). To have any plausibility at all as a treatment of rationality, this approach must have some way of explicating likelihood of truth that is not purely reliabilist. One way, perhaps not the only way, to meet this requirement is to say that the likelihood of truth of the judged content must be established a priori only from the possession-conditions of the concepts involved. (It may be tempting to add here "… and from information on which the thinker is entitled to rely in a relatively a priori fashion". But if it is rational to rely on that information, that itself is arguably a case of rationality that itself needs explication.) As any reader of the recent literature will be well aware, there are multiple issues to be resolved in elaborating such a position in detail. My main point here is simply to endorse a conception of the elucidation of the rationality of a

judgement that draws on connections between the fundamental reference-rules for the concepts in the judged content, and the fulfillment of the truth-conditions of that content.

There is a natural way under this conception of carrying out the second task, of explaining the thinker's appreciation of the location of the boundary between rational and non-rational judgements in given circumstances. The thinker herself draws on her tacit knowledge of the fundamental reference rules for the concepts in a given content *p*, and subpersonally draws on this information in coming to distinguish which contents are rationally judged in given circumstances, and which are not. We need to attribute tacit knowledge to explain a thinker's capacity to recognize the boundary between the rational and the non-rational just as we need to attribute tacit knowledge to explain the thinker's capacity to distinguish between grammatical and ungrammatical sentences of her own language. The nature of the need is the same in both cases. In both cases, there is an open-ended, unbounded to ability to classify correctly that must have some finitary basis.

This is barest outline of a position on the relations between understanding and rationality. But it is enough to permit us to draw some conclusions about the objectivity of judgement, and about disagreement in the special case of those concepts that are logical constants. I take the general issue of objectivity first.

The position I have outlined supports an entirely general case for its being an objective matter whether an arbitrary judgement is correct, a matter whose nature is constitutively independent of any dispositions to judgement that the thinker may have. Each content that can be judged is composed of concepts. Each such concept is individuated by its fundamental reference rule, the rule that states what makes something of the appropriate category the reference of the concept as employed by a given thinker at a given time in a given thinking. Such fundamental reference rules include the following, some of which I give by way of illustration, some of which I give to show how the relation-based concepts discussed earlier would be included in this framework:

> What makes an object fall within the extension of the observational concept *round* is that it is of the same shape as things are presented as being when they are perceived as round.

> What makes something the reference of the perceptual-demonstrative concept *that F* made available to a thinker by her perception at a given time in a given way W of something as F is that it is the F perceived then by her in way W.
>
> What makes something the reference of the first-person concept *I* as it occurs in a given thinking occurring in the thought of a given person is that the person is the author (the agent) of that thinking.
>
> What makes something fall within the extension of a thinker's concept *pain* is that there exists some subject of consciousness who experiences that event and for whom it has the same subjective character as events of the kind that she can recognize in herself as pains.

Each of these rules specifies a non-trivial condition for something to fall under the concept it treats. It is sometimes easy, and sometimes very difficult, to formulate the fundamental reference rule for a concept. But none of these rules is of a trivial disquotational form such as '*pain* is true of just the pains', or 'An arbitrary object fall under the concept *round* iff it is round'. Yet these particular rules illustrated do not offer reductions, or definitional eliminations, of the concepts they treat. We as theorists actually use the notions of being round, and of being in pain, in specifying the fundamental reference rules.

So, when a thinker judges a thought of the form Fa, the thought will be true if the object determined by the fundamental reference rule for the concept a, as used by the thinker in this judgement, falls within the extension determined by the fundamental reference rule for the concept F, as used by the thinker in the same judgement. An analogous point holds whatever the form of the content judged. There is a correctness-condition fixed by the fundamental reference rules themselves, as applied to the context of the thinker's judgement. Correctness of the judgement is not explained at all in terms of the thinker's dispositions to judge the content, however restricted or qualified these dispositions are. The position is, then, in strong contrast to those forms of linguistic

idealism, inspired by (possibly problematic) readings of the later Wittgenstein. According to linguistic idealism, the rule-following considerations require rejection of the idea that the correctness of a judgement depends on two factors: the nature of the content judged, and the facts of the situation; or, equivalently in the linguistic case, upon the meaning of the sentence asserted, and the facts of the situation.[8] By contrast, the account in terms of fundamental reference rules squarely endorses this classical two-factor conception of the truth of a judgement or an assertion. The account aims to give some theoretical underpinning for this intuitive conception. On this account, a species of minimal objectivity is built into judgement by the very individuation of the conceptual contents that compose the content judged.

This is entirely compatible with the existence of special cases in which judgements cannot, given the circumstances in which they are made, be mistaken. There are cases - 'I am in pain', 'This tastes sour' – in which if the judgement is made with understanding, on the basis of the circumstances that make it rational, then it will be true, for reasons related to the nature of the understanding and reasons in question. But this is not a consequence of any kind of general linguistic idealism. We have rather a class of special cases in which the understanding-based reasons for making the judgement are ones that ensure that the judgement will be true, given the fundamental reference-rules for the concepts involved.

This minimal objectivity is grounded in the nature of what is judged. It is formulated in terms of what is judged, and its correctness conditions, rather than in terms of warranted assertibility or in terms of the consequences of what is judged.[9]

The minimal objectivity of which I have been writing is very likely only the weakest kind in a spectrum of more demanding notions of objectivity. Particular domains

---

[8] Such an articulation of the later Wittgenstein's position is given by Crispin Wright, in his *Wittgenstein on the Foundations of Mathematics* (London: Duckworth, 1978).

[9] In this respect it differs from a notion of minimal truth that is explained in terms of warranted assertibility that is developed in Crispin Wright's *Truth and Objectivity* (Cambridge, Mass.: Harvard University Press, 1992), esp. Chs. 1, 2. For further comments on Wright's notion, see my *Truly Understood*, pp.43-45, or my paper 'Justification, Realism and the Past', *Mind* 114 (2005) 639-70, at p.662ff.. If what I have said earlier is correct, there are no substantive, understanding-based warranted assertibility-conditions for statements about other places, other times and other minds.

or subject-matters, such as the physical, or the mathematical, may enjoy stronger kinds of objectivity, kinds which apply specifically to the subject-matters mentioned in the reference-rules for the concepts of these specific domains.

Even minimal objectivity substantially constrains rational acceptance and rational dispute. If a content is true, there must be a non-trivial account in terms of the fundamental reference rules of why it is true. In this specific and limited sense, which is not committed to any reductionism about the subject-matter in question, the content cannot be 'barely true'. Correlatively, if a content is to be rationally held, the thinker must either have evidence that this non-trivial condition is fulfilled, or we must explain why she is rationally entitled not to require such evidence. Since understanding involves only tacit knowledge of these reference rules, and tacit knowledge need not at all be correctly articulated by its possessors, actual disputes over a content need not be formulated explicitly in terms of these fundamental reference rules. But rational acceptance, and rational disputes, must be construable in terms of these fundamental reference rules (or whatever the correct formulation of the fundamental reference rules may be) if they are to have the status of rational acceptance and rational disputes respectively.

What are the consequences of the outlined general conception of the relations between understanding and rationality for the case of logical concepts and logical principles?

Logical constants are special cases in at least two respects, when considered in the context of a theory of concepts as individuated by their fundamental reference rules.

First, the fundamental reference rules for logical constants do not mention any particular mental states such as perception, sensation or thinkings. In the case of an observational concept such as *round*, we can distinguish a part that concerns conditions for the concept to apply to an object that are given directly in terms of perception, and a part that applies a tacitly known condition, concerning sameness of shape, to this local, perceptual case. For logical concepts, there is no part of the first sort. There is only a tacitly known condition that is applied to the semantic values of the concepts of whatever category, or the propositions, on which the logical concept operates.

The second respect in which the logical concepts are arguably a special case is that the condition applied to the concepts or propositions on which the logical concept operates is one whose extension is determined a priori, given the extensions of what it operates on.[10]

I will mention three applications to logical concepts of the general position on concepts I have been outlining.

(1) Even in the somewhat rare cases of principles of which it really is required that a thinker must who grasps a particular logical concept must accept them if she is to possess the logical concept, such acceptance is rational, can be attained by thought, and is not simply the acceptance of a stipulation. Conjunction-introduction and conjunction-elimination are arguably such principles that must be accepted by someone who fully possesses the concept of conjunction. The rationality of accepting them is explained by the fundamental reference rule for conjunction that a thought of the form A&B is true iff A is true and B is true. The thinker's appreciation of the rationality is explained by the thinker's tacit knowledge of this fundamental reference rule. Her tacit knowledge of it explains the conditions under which she will evaluate a conjunction as true on the basis of information about the truth-values of its constituents. This in turn is a resource on which she will draw in coming to appreciate, rationally, that conjunction-introduction and conjunction-elimination are always truth-preserving. A thinker presented with these rules, or instances of them, explicitly for the first time is in a position to work out that they are correct for the concept of conjunction that she has been using for many years, and has fully understood for many years. The fact that we have difficulty in making sense of the possibility of someone fully understanding conjunction whilst not accepting these particular rules should not be used as evidence in favour of a conceptual-role theory of logical concepts.

(2) The tacit knowledge of a fundamental reference rule may have consequences for rationality that outrun what the thinker has hitherto appreciated. We can make sense of the idea that there are principles and axioms, even primitive principles and axioms,

[10] See my paper "What is a Logical Constant?", *Journal of Philosophy* **73** (1976) 221-240.

that are validated by the tacitly known fundamental reference rule, even though the thinker has not yet thought of them. Double-negation elimination was already valid for the concept of negation before anyone thought of its instances or of its general schema. When it is first accepted by a thinker for a concept of negation she already possesses, the rational acceptance is explained by her prior understanding, rather than that acceptance determining or contributing to what makes it the case that her understanding is of one concept rather than another. Understanding of the negation sign consists in the thinker's possession of the tacit knowledge that a sentence of the form '~A' is true iff A is not true.

(3) If a logical principle is rejected, or endorsed, and the rejection or the endorsement is to be correct, then there must be an underlying semantics on which the rejection or endorsement can be shown to be correct. (For the rejection or endorsement to be rational, it is required only that it be reasonable to think that there is such a semantics.) On the present treatment of understanding and its relation to correctness, this is a condition that holds both in the nonlogical cases, and in the special case of logical constants and their principles. This position is, then, one which definitely rejects Wittgenstein's own remarks, added for insertion into the *Philosophical Investigations*, in which he says:

> "There cannot be a question whether these or other rules are the correct ones for the use of "not". (I mean, whether they accord with its meaning.) For without these rules the word has as yet no meaning; and if we change the rules, it now has another meaning (or none), and in that case we may just as well change the word too."[11]

This general requirement for which I have been arguing, the requirement of a rationale for the acceptance and rejection of logical principles, is not specific or unique to referential semantics. If a thinker uses an expression for which the semantics are given by a species of probability, as on some treatment of the indicative conditional, the requirement still bites. The axioms and inferences rules in the resulting logic still have a

[11] *Philosophical Investigations*, tr. G. E. M. Anscombe, Third Edition (New York: Macmillan, 1968), p.147.

rationale in terms of the conditions on subjective probability, as for instance in Ernest Adams' semantics.[12]

Since, as I have repeatedly been saying, the ability to think about the concepts one is employing does not give one any privileged access to what individuates those concepts, arguments about the correct principles for logical and other concepts have to take the broad form of appeals to best explanation of a range of facts about reasons, judgements, and what is actually claimed in judging a proposition containing the concept in question. What I have said does not rule out revisions of all sorts in our thought, but it does constrain such revisions by the requirement that they be accompanied by a rationalizing semantics of a broad explanatory kind.

Nothing here is meant to advance the hopeless project of offering an entirely general noncircular justification of logic. Rather the position implies a constraint that must be fulfilled whatever our conception of the correct logic, if it is going to be possible to accept its principles rationally.

These points bear on the limits of intelligible disagreement more generally. If you are disputing a logical principle, or proposing a new one, you are committed to the existence of substantive semantics that explains why the principle you reject is a fallacy, or explains why the new principle you propose is valid under the semantics you endorse. The same applies to moral and to other areas of disagreement. We need an account of the moral, normative or other concepts involved which is such that tacit knowledge of that account explains why rejection of a previously accepted principle, or adoption of a new one, is correct. The account need not be reductive of normative concepts, and it may, though it need not, take the form of an analogue of the reference rules for those concepts in which some local, paradigm cases for application of the concept are picked out in a certain way, and application of an identity relation to those cases determines the extension of the concept elsewhere. The substantive requirement is just that there be some such validating account, whatever its form.

So much for some of the consequences of this conception of understanding and rationality for the case of logical concepts. The account also raises further issues, most

[12] E. Adams, *The Logic of Conditionals* (Dordrecht: Reidel, 1975).

immediately: what is it for a subpersonal state of tacit knowledge to contain a particular logical notion, for example, the notion of alternation or negation, in its content? We make such attributions when we say that to grasp alternation is to have tacit knowledge that something of the form 'A v B' is true iff either A is true or B is true. One untempting answer is that it is to have some expression for the logical notion in a language of thought, and also to possess some complete set of rules, or of axioms and inferential principles, for alternation at the subpersonal level, where this set is drawn upon in subpersonal computations. There are at least two (and maybe more) reasons this is unattractive. First, if all the rules and axioms are given equal status, this apparatus is not clearly explaining why some primitive rules for negation have to be worked out on the basis of our understanding in a way that is more demanding for us than, say, the introduction and elimination rules for conjunction. Second, this style of approach is in any case not available for more sophisticated logical or mathematical notions where the axioms or rules are not recursively enumerable, but where we nevertheless still have a grasp of a determinate concept.

A much more attractive general style of treatment, which could be developed in a number of different detailed ways, is to say that a thinker's having tacit knowledge of some principle such as

'~A' is true iff A is not true

consists in his evaluating mental models, in which truth-values are assigned to propositions or contents, in a certain way when the propositions involve negation.[13] This too is not meant to be an eliminative or reductive account of what it is to grasp negation (a hopeless and unnecessary task), but rather a constraint upon it. Some operations on models may be found harder by a thinker, some may be found easier, and this can explain differences in the ease of appreciation at the personal level of which axioms and principles are correct, in that they hold in all models of a certain kind. Introducing more model-theoretic material into our computational procedures also promises an account of what is involved in our grasp of notions for whose principles there is no recursively

[13] For a general conception of mental models, see P. Johnson-Laird, *Mental Models* (Cambridge, Mass.: Harvard, 1983). Not all of his theories about how these models operate are essential to any treatment of mental models.

enumerable axiomatization. Our notion of correctness in such cases cannot simply be that of following from a certain recursively enumerable set of axioms. It is more plausibly that of something that is true in all models of a certain kind.

The fundamental reference rules I suggested for atomic predicates and singular concepts are not of the trivial, disquotational form. Yet my overall statement of the content of the tacit knowledge involved in understanding a classical logical constant is of disquotational form. Why should the form be acceptable for a logical constant, but not for atomic predicates and singular concepts? I said that for the predicates and singular concepts, the non-disquotational form is needed to exhibit the finitary basis of the thinker's appreciation of the circumstances in which contents containing the concept may be rationally accepted, rationally rejected, or rationally held open. With the logical concepts, however, once the thinker has a grasp of rational conditions of acceptance and rejection for atomic propositions, an understanding – even a disquotational understanding – can be a basis for rational acceptance or rational rejection of complex contents built up by logical operators from contents already grasped. This is connected founded in the fact that the semantic value of a thought or sentence whose principal operator is a logical constant is a priori determined by the semantic values of the operands, together with the semantic rule for the logical operator in question.

Almost everything I have said in this paper relies on the existence of explanatory states that have subpersonal characterizations, possess content, and explain content-involving phenomena, including the phenomena of conscious thought and reason. In certain domains, including early vision, and also the perception of syntactic structure, it would be almost universally accepted by cognitive scientists that content-involving states of tacit knowledge have a real role in explaining personal-level content-involving phenomena. There are pertinent questions, both empirical and conceptual, about how such explanations operate. By now, however, these questions seem much more issues about how and why a successful enterprise is successful, rather than questions which threaten to undermine that whole conception of explanation in these domains.

In the domain of semantics, it would not be true to say that such tacit knowledge has been wholly unexploited. It has in fact been exploited in proposed explanations of the

perception of sentences, phrases and words as having particular meanings. But many of these semantic applications have essentially been to the phenomenon of the association of particular meanings with particular sentences, phrases, and expressions in those sentences. To use Dummett's terminology, the applications in semantics have been to phenomena at the level of meaning-theories for particular languages and speakers' grasp of those meaning-theories, rather than at the level of the theory of meaning in general.[14] Part of the burden of this paper has been that the explanatory powers of states of tacit knowledge should be applied also to the unavoidable question of what it is to have the concepts that are associated with particular expressions in a language – not only to logical concepts, but to concepts in general. Unless we do so, we will not have an adequate explanation of the full range of the phenomena of rationality, understanding and the limits of intelligible disagreement.

Columbia University in the City of
New York
University College, London

[14] M. Dummett, *The Logical Basis of Metaphysics* (Cambridge, Mass.: Harvard University Press, 1991), p.22.