

Conceptual Analysis, Theory Construction, and Conceptual Elucidation

Julia Tanney, Kent

Almost a half century after the publication of the *Philosophical Investigations*, it seems important to ask why Wittgenstein's ideas have had so little impact on contemporary discussions in the philosophy of mind.

A clue can be discerned by what Georges Rey says in the introduction to his book on contemporary philosophy of mind. Rey announces at the outset to his readers that his treatment of the mind aspires to be continuous with science, not with literature. He explains that there is a recent resurgence of interest in the philosophy of mind with "explanatory questions" about what sort of thing a pain, a thought, a mental image, a desire, or an emotion is. Neither materialism nor dualism provides a "serious" theory about the mind, which will give us a "serious" explanation of mental phenomena. According to Rey, although old-style grammatical investigations may have given us a "heightened sensitivity to complexities and nuances of our ordinary mental talk," they "tended to occur at the expense of further theorizing about the mental phenomena themselves" (Rey, 4).

But, as Rey acknowledges later in the book, it is not simply that the acquisition of a heightened sensitivity to the complexes and nuances of ordinary talk *tended* to occur at the expense of further theorizing about the nature of mental phenomena: it was *designed* to do so. Such a "heightened sensitivity" was to reveal the futility of such further theorizing. Wittgenstein explicitly repudiates the temptation to *penetrate* or *see into* phenomena. He is also complains about attempts at explanation that invoke "yet uncomprehended processes in yet unexplored mediums." As philosophers, we should be interested in the kind of statements we make in using mental concepts.

Comments like this have led many contemporary philosophers to reject Wittgenstein's attitude as simply prejudiced against scientific enquiry. Rey parodies this attitude – crystallized in Wittgenstein's dictum that explanations have to come to an end somewhere – by suggesting that

one could... [rationally]... never ask for explanations of anything at all: we *could* just say that it is a natural capacity of lightning to burn what it strikes. The question is whether we *can also* rationally ask for slightly deeper explanations, and, if we can, what those explanations might be... (p.5).

But Wittgenstein does not simply reject scientific theories that are advanced about the phenomena in question: he disputes whether there are such phenomena to be so investigated. He rejects, that is, the philosophical presupposition that funds the scientific programme: namely, that expressions containing mental verbs always function in the same way to pick out an independent state of affairs (objects, events, states, properties, or relations) in the first place. This is not anti-science; it is anti-Russell and anti-Tractatus. Expressions containing mental verbs function in various ways at various times. To suppose that they always refer, let alone to phenomena that are ripe for metaphysical or scientific investigation, would be unacceptably to distort the meaning of these expressions. This is not because concept revision is in principle

unacceptable, but because these particular kinds of mental expressions, and the kinds of explanations that they deliver, are perfectly in order as they are. This becomes clear when close attention is paid to the way these expressions are used. It is a great pity that today this fundamental tack is rarely discussed or even acknowledged.

But an interesting trend is worth noting. Cognitive science's close relationship with analytical philosophy is being questioned, and as this occurs, some of Wittgenstein's thoughts are getting a look-in through the back door. This can be seen with respect to a project that has occupied most philosophers of mind in the 1980s and 90s: to articulate a theory of mental content.

Stephen Stich characterises one motivation for articulating such a theory as follows:

A prominent feature of our everyday discourse about ourselves and about other people is our practice of identifying mental states by adverting to their content... In these, and in a vast range of other cases, the attribution of content is effortless, unproblematic, and unquestionably useful. Moreover, in the typical case, there is widespread intersubjective agreement about these attributions. Plainly, there must be a mental mechanism of some complexity underlying this ubiquitous practice, and it seems plausible to suppose that the mechanism in question includes a store of largely tacit knowledge about the conditions under which it is (and is not) appropriate to characterize a mental state as the belief or the desire *that p* (p. 350).

Stich suggests that a perfectly plausible goal for a theory of mental content would be to describe the body of tacit knowledge that is stored in the mechanism underlying our quotidian practice.

If a theory of content is to explain how an ordinary person comes to an understanding of her subject by attributions of mental states then the theory is answerable to the kind of understanding that is involved, and thus the kind of explanation that is delivered, in common sense interpretive contexts. Even if a theorist studying these practices is the best person to articulate what constitute their rules or suppositions, it is still the ordinary practices, explanations, and understanding they deliver that is the subject of investigation.

But philosophers working on theories of mental content have a highly ambivalent attitude toward ordinary understanding. Often, one is given the impression that the authority on the nature of the explanatory role of the mental is within the province of the cognitive sciences, whose theorists should not feel constrained by the remarks of philosophers interested in conceptual elucidation. Anybody tempted to appeal to ordinary language "intuitions" is interfering unnecessarily with a perfectly legitimate scientific enquiry.

The relationship was bound to become uneasy between the tradition of "analytical" philosophy – which takes as its brief an investigation of language as a way of investigating thought, and philosophy-as-science – which

wants to eschew language (especially with its tight ties to what the ordinary folk would say) and study the phenomena of thought on its own. One manifestation of this uneasiness is the realization that theory construction is not far removed from the project of conceptual analysis as practiced, for example, by G.E. Moore.

Moore's explicit characterisation of analysis – as the clarification of concepts – says that the verbal expression of analysis must follow a standard pattern of paraphrase in which what is analysed is logically equivalent to a larger, more explicit, and synonymous expression. Of course the accounts of the mental we are concerned with were never supposed to be definitional, hence the demand for synonymy of expressions is no longer in place. Instead, the strategy might be to suppose that our mental concepts fix the reference of what can be discovered *a posteriori* to be some property or set of properties. On this conception, any proposed identity claim linking the concept with the alleged referent will not purport to be definitional, but rather “reforming”. Nonetheless, any theoretical or reforming definition might still be subject to the challenge that it has either left something essential out of the proposed reforming definition that is central to the concept being reformed, or that it has added something that it was essentially missing.

A number of philosophers have noticed the close tie between theory construction that proceeds in this way and old-fashioned conceptual analysis and have thus expressed misgivings about the whole enterprise. Stephen Stich and Michael Tye have doubts partly on inductive grounds: since conceptual analysis or conceptual definition as Moore conceived it has, by and large, been a failure, then so too will reforming definitions with (in certain ways) even tighter constraints about what is allowed in the paraphrase.

But the past failures of conceptual analysis are not the only grounds for pessimism. Both Stich and Tye have also claimed that empirical findings support rival theories about the nature of the mental structures that underlie people's judgments when they classify items into categories. The empirical evidence shows that these structures do not exploit tacitly known necessary and sufficient conditions. Both philosophers cite Eleanor Rosch's work on prototypes as an example of such a rival theory and both of them mention in a footnote her indebtedness to Wittgenstein's work in the *Philosophical Investigations*. This is what I mean by a recent attention to Wittgenstein slipping through the back door.

What conclusion is to be drawn from this? Stich speculatively concludes that the ability to categorize intentional concepts will not have underlying them mental mechanisms that utilize “classical” concepts of the sort that can be defined by sets of necessary and sufficient conditions. If our intuitions about whether a state has the content *that p* are constructed, and not determined by rules for the application of the concept but are rather guided in part in response to the circumstances in which the judgment is called for, then there are no such rules in the form of necessary and sufficient conditions for a theory of mental content to deliver. To a student of Wittgenstein (and ignoring, for the moment, the reference to mental mechanisms), this will sound familiar.

One would expect a concomitant rejection of the idea that it is the job of cognitive science to describe the structure of knowledge underlying people's ordinary judgments about the content of intentional states. But instead Stich suggests that if this is what a theory of mental content is for, the theorist will probably have to give

up doing philosophy as it is traditionally conceived and start doing cognitive science instead (p.354).

But how would doing cognitive science instead avoid what has always been the difficult questions that a theory of content would have to face: namely, what is to count as someone's having a particular mental attitude with a specific content? Or, as Wittgenstein might prefer to say, what counts as a successful or permissible or acceptable use of these mental expressions? Experimenters testing “underlying mechanisms” for their subjects' ability to identify these states or attribute mental concepts have to take a stand on whether any particular attribution is successful before the mechanisms for delivering these and only these successful judgments can be investigated. To reach the conclusion that concepts are not classical in the sense that Moore envisaged does not obviate the need to find out or decide what will count as a successful application of these concepts. Insofar as it is a person's interpretive *abilities* to be explained (and not mere habits or tendencies) we need some idea of what constitutes success.

The pursuit of these answers was traditionally the remit of philosophy. But in a context where the focus is on the mental mechanisms in which the relevant bodies of knowledge are held to be stored, the role of the philosopher has become tenuous. For is it not the business of an empirical psychologist to study mental mechanisms? Thus both Stich and Tye agree that whatever the mental is, it is the job of the latter to investigate it.

Stich and Tye may be right to doubt that a body of knowledge provides necessary and sufficient conditions (for something, say, to be a state with content) but they should – for similar reasons – also be suspicious of the idea that our ability to apply mental concepts has a body of largely tacit knowledge that “underlies” it. How is this knowledge supposed to be explanatory?

The usual supposition is that an interpreter is caused to deliver correct interpretations most the time because the rules for interpreting are embodied in “mental mechanisms.” This presumably means that having the body of knowledge is causally sufficient to produce the correct answers, provided that the mechanism is not defective. But if the mechanism that embodies the knowledge is causally sufficient to produce the right answers, then the content of knowledge itself, as it were, would have to suffice to deliver the right answers. And this leads us back to the idea that the knowledge consists in at least sufficient conditions for the application of our content concepts. Surely the rejection of a classical conception of concepts brings with it the rejection of the model of explanation that supposes that the content of knowledge is sufficient to determine correct interpretations. To put the point another way, surely the rejection of the idea that there are rules that determine the correct application of mental concepts brings with it the rejection of the idea that our ability to interpret is a matter of knowing these rules. What room is left, then, for the idea that the ability to be explained amounts to a kind of information processing? How is the information to be delivered and how is its delivery supposed to explain the ability?

It seems that in viewing a theory of content as an underlying body of knowledge that is supposed to explain an individual's ability to interpret, we have not moved far enough away from conceptual analysis. Moving far enough away would involve rejecting the mechanism hypothesis as otiose. This will in turn allow the focus to be on the intersubjective agreement: why do we tend to converge on the ascription of mental content concepts to one and

another, and in what do correct or permissible judgments consist? These questions will be answered – to the extent that they can be – by paying attention to how we justify, defend, correct, and explain those attributions. And this is just what a good Wittgensteinian should set out to do: to describe the “allowable moves” that can be made using expressions containing mental verbs. Insofar as the focus of investigation is on what counts as “allowable” – as opposed to what is known by someone who makes the moves – the project would be normative in the old-fashioned sense, but it is as far removed from the presuppositions of *a priori* conceptual analysis as the critics should want.

References

- Rey, G. 1997 *Contemporary Philosophy of Mind*, Oxford: Blackwells.
- Stich, S. 1992 “What Is a Theory of Mental Representation?”, reprinted in 1994 Stich and Warfield (eds.) *Mental Representation*, Oxford: Blackwells, 347-364.
- Tye, M. 1992 “Naturalism and the Mental”, *Mind*, 101, 421-41.