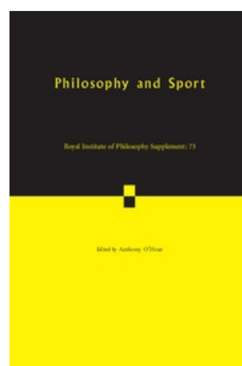
Royal Institute of Philosophy Supplement

http://journals.cambridge.org/PHS

Additional services for ***Royal Institute of Philosophy Supplement:***

Email alerts: Click here
Subscriptions: Click here
Commercial reprints: Click here
Terms of use : Click here



Self-knowledge, Normativity, and Construction

Julia Tanney

Royal Institute of Philosophy Supplement / Volume 51 / March 2002, pp 37 - 55
DOI: 10.1017/S1358246100008079, Published online: 12 April 2010

Link to this article: http://journals.cambridge.org/abstract_S1358246100008079

How to cite this article:
Julia Tanney (2002). Self-knowledge, Normativity, and Construction. Royal Institute of Philosophy Supplement, 51, pp 37-55 doi:10.1017/S1358246100008079

Request Permissions : Click here

CAMBRIDGE JOURNALS

Downloaded from http://journals.cambridge.org/PHS, IP address: 129.12.11.80 on 10 Nov 2013

Self-knowledge, Normativity, and Construction

JULIA TANNEY

He tried to look into her face, to find out what she thought, but she was smelling the lilac and the lilies of the valley and did not know herself what she was thinking—what she ought to say or do.
Oblomov

1. Much of modern and contemporary philosophy of mind in the 'analytic' tradition has presupposed, since Descartes, what might be called a realist view about the mind and the mental. According to this view there are independently existing, determinate items (states, events, dispositions or relations) that are the truth-conferrers of our ascriptions of mental predicates.[1] The view is also a cognitivist one insofar as it holds that when we correctly ascribe such a predicate to an individual the correctness consists in the discovery of a determinate fact of the matter about the state the individual is in—a state which is somehow cognized by the ascriber. Disputes have arisen about the nature of the truth-conferrers (e.g., whether they are physical or not) and about the status and the nature of the individual's own authority about the state he is in. A dissenting position in philosophy of mind would have to be handled carefully. It would, most importantly, need to allow for the objectivity of ascriptions of mental predicates at least insofar as it made sense to reject some and accept others on appropriate grounds. Perhaps such a position in the philosophy of mind can be likened in at least one way to what David Wiggins has characterized as a doctrine of 'cognitive underdetermination' about moral or practical judgments.[2] In comparing his position of cognitive underdetermination about moral or practical judgments to some things Wittgenstein has said about the philosophy of mathematics, Wiggins suggests that, 'In the

[1] Henceforth, I shall speak of *states* or *events* for ease of exposition. By 'independently (or antecedently) existing', I mean states whose existence does not depend on any epistemic interest the subject might take in them.

[2] See 'Truth, Invention and the Meaning of Life', 'A Sensible Subjectivism', and 'Truth, and Truth as Predicated of Moral Judgements', reprinted as essays III, IV, and V in *Needs, Value and Truth* (Oxford: Basil Blackwell, 1987).

assertibility (or truth) of mathematical statements we see what perhaps we can never see in the assertibility of empirical (such as geographical or historical) statements: the compossibility of objectivity, discovery, *and* invention.'[3]

In this paper I intend to develop the idea that the 'compossibility of objectivity, discovery and invention' is a part of our ordinary (i.e., non-scientific and non-theoretical) understanding of the mental.[4] If this is correct, it is important, since contemporary theories do not make sense of this compossibility: they fail, in particular, to leave room for the inventive aspects of self-ascription.[5]

My strategy involves appealing to intuitions about the acceptability or appropriateness of certain ascriptions of mental concepts. To generate these intuitions it will be helpful to rely on the description of a possible person and her thoughts made out both in a certain degree of detail and over a significant period of time. I shall be looking at selected details from Goncharov's *Oblomov* and asking my reader to consider some of the scenes concerning the character Olga.[6]

2. The so-called 'Cartesian' model of the mind is a model in which discovery plays a role *par excellence*. This model supposes that one's thoughts, feelings, concerns, needs, values and principles, are played out in an 'inner theatre' of the mind that is constitutively independent of any epistemic interest the subject might take in it. More problematically, it also supposes that these items are available completely and unmistakably as a result of introspection.

Consider carefully what this picture is committed to. Our experience of these events is thought to be unmediated in the sense that it would not be subject to norms or rules; these events would be

[3] 'Truth, Invention and the Meaning of Life', op. cit., p. 130.

[4] This paper develops some of the ideas I suggested in 'A Constructivist Picture of Self-Knowledge' *Philosophy*, **71**, no. 277 (July 1996), pp. 405–22.

[5] That they cannot make sense of it comes as no surprise, once it is noted that the goal of so-called 'naturalism'—to locate the mental within nature conceived as the realm of law—is *ipso facto* removing from the mental the first-personal point of view or participant perspective that seems so important for retaining the inventive aspect. For this reason it would seem as if any theory of mind that conceives its starting point as the recoil from dualism—(e.g. behaviourism, identity theories, functionalism, and even anomalous monism)—and attempts a full-bodied or modified physicalism, will be unable to account for the inventive or constructive aspect of the mental.

[6] All references are to the Penguin edition, translated by David Magarshack, 1954.

simply *given* as part of our immediate experience. If our ability to classify them and to recognize them, as classified, is to be infallible this means that our bringing them under concepts—and this would be a matter of applying rules—would not be subject to error. Introducing names for these experiences into our language would presumably involve simple association of the object or event with a name, or 'baptism' by ostensive definition. Such classifying or naming, however, would be a private activity, since the experiences classified and named are not accessible to others and the associations cannot be checked by anyone else.

There is a problem with this view. The Cartesian wants the inhabitants of the mind to have an existence that is independent of any epistemic interest taken in them; in this sense he is a realist. He must therefore allow a sufficient gap between what is grasped when the subject 'turns his mental eye inward' and his grasping it. One might reasonably require that in order to effect this gap and bring out the true independence of the nature of the objects of the mind it has to be *in some sense* possible for the subject to get it wrong. But the introduction of infallible access thwarts this possibility.

The Cartesian might attempt to dig in his heels and claim that the fallibility associated with our sense-perception of the external world simply fails to apply to the perception of our own minds. Whereas in sense perception the possibility of error is a mark of the independence of the object perceived, he might deny that a viable realism about the mental requires such a possibility. It just requires that there be a mental item or state that is constitutively independent of the subject's gaze. Success is assured, then, since on this view the 'mental eye' and its conceptual machinery functions perfectly.[7]

The opponent of this view must then turn his attention to the idea of perfectly functioning conceptual/perceptual equipment. And, indeed, he might plausibly maintain that the very idea of an explanation that posits perfectly functioning machinery is of dubious coherence. For if we posit a mechanism that functions perfectly and cannot go wrong, then we cannot appeal to this mechanism as an *explanation* of the ability. The reason is simply that there would be no way to distinguish any purported explanation using a mechanism that cannot go wrong from a mere description of what would constitute success. As long as it is explanatorily indistinguishable from such a description, there is no reason to posit the mechanism to begin with. And if there is no perfectly function-

[7] See Crispin Wright's discussion of the Cartesian view in 'Wittgenstein's Later Philosophy of Mind: Sensation, Privacy, Intention', *Meaning Scepticism*, K. Puhl (ed.) (Berlin: de Gruyter, 1991), pp. 126–47.

ing mechanism, then the whole idea of objects before the mind that are perceivable by this mechanism is threatened.

The Cartesian model invites us, in effect, to compare the referents of our mental concepts 'in their definiteness to objects which are already lying in a drawer and which we then take out.'[8] To give this realist aspect of the Cartesian view—the idea that mental items exist in their definiteness independently of any act of identification or endorsement—more chance of success, let us disentangle it from the Cartesian notion of infallibility. The idea that a person might not be aware of what he is thinking is an idea that many people nowadays will be happy to accept. (Many feel this was a discovery of Freud; an idea that along with the Freudian notion of the 'unconscious' has not only permeated our commonsense psychological practices but, in the kindred (though in aspects quite different) form of 'tacit' knowledge, has permeated contemporary theorizing about language and mind in the cognitive sciences. Some of the minority who remain sceptical about unconscious thoughts have even indicated that the fallibility of the first-person ascriptions stands or falls with Freud's technical notion of the unconscious. In my view, both ideas are wrong. Freud's examples of *parapraxes* in *The Psychopathology of Everyday Life* were convincing because he identified, and put a name to, patterns of action and speech that were candidates for motivated behaviour that could be recognized as such by anyone to whom the patterns were pointed out (including the agent himself). I am not familiar enough with literary texts to know when authors started exploring the idea that the intentions and motivations could be discerned without the agent's awareness. But the idea figures commonly in Russia in the works of Dostoyevsky and Goncharov (the latter began writing *Oblomov* in 1849). In France it is evidently to be found in the work of Diderot (*Jacques le Fataliste* (written in 1773)) and it is a major theme in Constant's *Adolphe* (1816), thereby predating the popularization of Freud's work at least in France by a century or more.)

Consider a scene from *Oblomov* in which the fallibility or, in this case, the incompleteness of the subject's own gaze is manifest. Here, the idea that the contents of Olga's mind are apt for 'discovery' is especially appropriate. The reader is made aware not only of Olga's

[8] The quotation is from §193 of Wittgenstein's *Philosophical Investigations*, trans. G. E. M. Anscombe (Oxford: Basil Blackwell, 1953). He uses the metaphor to illuminate the idea that a machine's action seems to be in it from the start (and the metaphor of a machine had been introduced in an attempt to make sense of the idea that an act of meaning can in some sense anticipate reality (§188)).

words but also of her own thoughts in the form of 'inner speech'. But some of her thoughts and feelings she is not yet able to recognize: that she is in love with Oblomov, that she was pleased (albeit flustered) by Oblomov's sudden declaration of love, and that she is horrified as he attempts to take it back. Oblomov speaks first, trying to make up with her after rashly declaring his love:

> 'Please believe me, the whole thing—I mean, I don't know what made me say it—I couldn't help it,' he began gradually growing bolder. 'I'd have said it if a thunderbolt had struck me or a stone had crashed on top of me. Nothing in the world could have stopped me. Please, please don't think that I wanted—I'd have given anything a moment later to take back the rash word. ...'
>
> She walked with her head bowed, sniffing the flowers.
>
> 'Please forget it,' he went on, 'forget it, particularly as it wasn't true....'
>
> 'Not true?' she suddenly repeated, drawing herself up and dropping the flowers.
>
> Her eyes opened wide and flashed with surprise.
>
> 'How do you mean—not true?' she repeated.
>
> 'I mean—well—for God's sake don't be angry with me and forget it. Please, believe me, I was just carried away for a moment—because of the music.'
>
> 'Only because of the music?'
>
> She turned pale and her eyes grew dim.
>
> 'Well,' she thought, 'everything's all right now. He took back his rash words and there's no need for me to be angry any more! That's excellent—now I needn't worry any more. ... We can talk and joke as before.'
>
> She broke off a twig from a tree absent-mindedly, bit off a leaf, and then at once threw down the twig and the leaf on the path. 'You're not angry with me, are you? You have forgotten, haven't you?' Oblomov said, bending forward to her.
>
> 'What was that? What did you ask?' she said nervously, almost with vexation, turning away from him. 'I've forgotten everything—I've such a bad memory!'
>
> He fell silent and did not know what to do. He saw her sudden vexation but did not see the cause of it.
>
> 'Goodness,' she thought, 'now everything is all right again. It's just as if that scene had never taken place, thank heaven! Well, all the better. ... Oh dear, what does it all mean? [...]
>
> I'm going home,' she said suddenly, quickening her steps and turning into another avenue.

> There was a lump in her throat. She was afraid she might cry. (pp. 207, 208)

Oblomov sees that his attempt to take his rash words back is more distressing to Olga than his original declaration of love but does not understand why. She is mollified when he is forced virtually to redeclare his love and he is left feeling confused. Only later, in reflecting on the moment when she breaks the lilac sprig, does he come to realize that she loves him. He approaches her again, carrying the lilac sprig, armed with this new knowledge.

> 'What have you got there?'
> 'A twig.'
> 'What sort of twig?'
> 'As you see: it's lilac.'
> 'Where did you get it? There is no lilac here. Which way did you come?'
> 'It's the same sprig you plucked and threw away.'
> 'Why did you pick it up?'
> 'Oh, I don't know. I suppose I was glad that—that you threw it away in vexation.'
> 'You're glad I was vexed! That's something new. Why?'
> 'I won't tell you!'
> 'Please, do, I beg you.'
> 'Never! Not for anything in the world!'
> 'I implore you!'
> He shook his head.
> [...]
> 'What's the matter? Is it something dreadful?' she said, her whole mind concentrated on the question, glancing searchingly at him.
> Then gradually realization came to her: the ray of thought and surmise spread to every feature of her face and, suddenly, her whole face lit up with the consciousness of the truth. ...Just like the sun which, emerging from behind a cloud, sometimes first lights up one bush, then another, then the roof of a house and, suddenly, floods a whole landscape with light. She knew what Oblomov's thought was.
> 'No, no,' Oblomov kept repeating. 'I could never say it. It's no use your asking.'
> 'I'm not asking you,' she replied indifferently.
> 'Aren't you? But just now —'
> 'Let's go home,' she said seriously, without listening to him. 'Auntie is waiting.' (pp. 216, 217)

After this scene, Olga goes home and immediately begins acting like a woman in love; I shall discuss this transformation shortly. For the present all we need notice is that these passages illustrate nicely the sense in which 'discovery' is an apt description of what sometimes happens in self- and other-ascriptions. It also illustrates nicely the sense in which these ascriptions might qualify as 'objective.' Whatever pattern of thought and behaviour is supposed to indicate a person's mental states, it is often identifiable by others. In this case Oblomov is the first to identify some of Olga's thoughts and feelings. When Olga finally comes to see them, her recognition results from inference or a chain of reasoning: in this case via her realization of Oblomov's thoughts about the significance of her behaviour.

3. One of the problems with the Cartesian model is its failure to leave room for a requisite sense of objectivity. Another is its failure to accommodate the intuition that at least for many mental states (paradigmatically ones involving 'propositional attitude' concepts) the criticism we incur when we misascribe results from a kind of explanatory failure. Indeed, what generally defeats a self-ascription is its failure to fit into a rationalizing story. According to the perceptual model, defeat is rather a matter of failing to track the private items that exist in the mind's eye. Of course, even on this Cartesian model what is tracked may be—contingently—(part of) an explanatory project. But later philosophy of mind has accorded mental concepts (especially those apt to play a role in reason-explanation) with more than a merely contingent explanatory role: the intuition—which forced those attracted to physicalism to withdraw to token physicalism—is that propositional attitude concepts in particular (and hence the emotional states that presuppose them) are —constitutively—explanatory concepts.

Functionalism—the dominant position in contemporary philosophy of mind—seems to avoid both objections to the Cartesian model and yet retain the sense in which self-knowledge is comfortably seen as a matter of discovery. This theory of the mental says that when we ascribe a mental concept to an individual, this concept refers to a state the person is in that has appropriate causal connections to sensory input, behavioural output, and other internal states.

According to this doctrine, success in ascription would fundamentally be a matter of tracking or homing in upon those states with the appropriate causal specification. Functionalists might still maintain that the point of mental-concept ascription is to render intelligible the one to whom the concepts are ascribed, as long as it is the causal role that is doing the explanatory work. This picture would

seem to be consistent with ideas exemplified in the scenes from *Oblomov* at least insofar as ascription is tantamount to the discovery of a pattern. It is also consistent with the idea that a person might not be in the best position to notice this pattern (as causal role).

Problems arise, however, when one reflects on what exactly is doing the explanatory work. Although I will not argue for this here, I think it is doubtful whether functionalists can consistently maintain that the explanatory project is a *rationalizing* project or that mental concepts are explanatory in virtue of the way in which 'things are made intelligible by being revealed to be, or to approximate to being, as they rationally ought to be'.[9] They will not be able to do this, at least, if such a style of explanation is 'to be contrasted with a style of explanation in which one makes things intelligible by representing their coming into being as a particular instance of how things generally tend to happen.'[10]

If, as I believe, these styles of explanation are indeed different,[11] then a question arises about functionalism's relation to our ordinary, commonsense ascriptive practices where these are understood as making fundamental use of *rationalizing* explanations. Functionalism, as originally conceived, was a thesis about the meaning of mental predicates—presumably about the meaning of those predicates used within our ordinary, commonsense, psychological practices. As such, its viability as a theory of mind would depend upon whether the concepts ascribed within these practices do, in effect, track internal, functionally individuated (physically realized), causally efficacious states or events.

My suspicion is that they do not. I mention this difference in explanatory patterns—between causal-explanatory patterns on the one hand, and rationalizing patterns on the other—because I suspect it will be of utmost importance. For it is arguably the rationalizing pattern and not (or not merely) the causal-explanatory one that allows the reintroduction of an inventive aspect to the role of self-ascriptions.

4. I now intend to consider what is intuitively attractive about the idea that a person has some inventive or creative role to play in

[9] The phrase is John McDowell's in 'Functionalism and Anomalous Monism', *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, Lepore and McLaughlin (eds) (Oxford: Basil Blackwell, 1985), p. 389.

[10] Ibid. p. 389

[11] See my 'Why Reasons May Not be Causes' *Mind & Language*, **10**, nos. 1/2, pp. 103–126.

respect to his mental life. I have discussed how it makes sense to say that Olga's realization of her love for Oblomov came as a discovery (one that Oblomov had made before her). But to call it a 'discovery' would only be partly correct. A few hours after Olga realizes her feeling for Oblomov, she becomes transformed.

> He waited nervously and with trepidation for Olga to come down to dinner, wondering what she would say, how she would speak, and how she would look at him. ...
>
> She came down—and he could not help admiring her; he hardly recognized her. Her face was different, even her voice was not the same. The young, naïve, almost childish smile not once appeared on her lips; she did not once look at him with wide-open eyes questioningly or puzzled or with good-natured curiosity, as though she had nothing more to ask, find out, or be surprised at. Her eyes did not follow him as before. She looked at him as though she had studied him thoroughly, and, finally, as though he were nothing to her, no more than the baron—in short, he felt as though he had not seen her for a whole year during which she had grown into a woman. (pp. 222, 223)

Olga is transformed from someone who was (arguably) in love into someone who now acts in self-conscious awareness of her love or in accordance with her own conception of how a woman in love should act. Might not this passage suggest that there is something right about the idea that the nature of the love she has 'discovered' is changed as a result of these subsequent actions, and hence as a result of this self-awareness? The idea would be not merely that her love for Oblomov causes her transformation into a 'woman' or even that her *awareness* of it does, but rather that her awareness and her endorsement of it somehow affect the love or the shape of the love itself. They play a role in a more complex 'rationalizing project' that involves her own conception of how a woman in love should act. This explanation of her behaviour (that she is in love with Oblomov), its endorsement by Olga, and its role in an ongoing rationalizing project give shape to, or articulate, a pattern or a possibility which in turn (retrospectively, as it were) supports the original explanation that Olga is in love.

Charles Taylor has, in a series of articles, attempted to argue for the idea (which he credits to Heidegger) that a person's self-conception partly constitutes the mental state he is in.[12] He argues that much of what we think, feel, and value is not the result of our being

moved by forces like gravity or electro-magnetism. Our desires and aspirations are given formulation in words or images; they cannot but be articulated or interpreted by us somehow. But these articulations

> are not simply descriptions, if we mean by this characterizations of a fully independent object, that is, an object which is altered neither in what it is, nor in the degree or manner of its evidence to us by the description. In this way my characterization of this table as brown, or this line of mountains as jagged, is a simple description.
>
> On the contrary, articulations are attempts to formulate what is initially inchoate, or confused, or badly formulated. But this kind of formulation or reformulation does not leave its object unchanged. To give a certain articulation is to shape our sense of what we desire or what we hold important in a certain way.[13]

Taylor gives his own example of what he means but the point can be developed by staying with the character of Olga and by noting how Olga's endorsement of herself as a woman in love gives shape to, or articulates what—although it amounted to a 'discovery'—had been inchoate or confused before.

An analogy might be helpful. Think about a duck-rabbit design, which, although ambiguous between being either the head of a duck or a head of a rabbit, is arguably not the head of a cow or pig. Now, imagine that when the figure is drawn with more detail (a body is added) it becomes a duck and not a rabbit. The analogy would be that Olga's pattern of behaviour before her reflections was in certain ways indeterminate (although certain interpretations of her behaviour could be ruled out) just as the duck-rabbit design is indeterminate or ambiguous (though certain interpretations can be ruled out). After her reflections and her endorsement of one pattern (she recognizes it as a duck), she behaves in a way that is consistent with that recognition. Her endorsement of it (as a duck) and her subsequent behaviour allow the pattern to develop in such a way (say, it develops a beak, webbed feet, feathers, etc.) that renders the other interpretation no longer viable.

This idea can be spelled out in more detail when we consider what happens to Olga later in the novel. Her relation with Oblomov has

[12] See especially 'What is Human Agency?', 'Self-Interpreting Animals', and 'The Concept of a Person', reprinted as chapters 1, 2, and 4, respectively in *Human Agency and Language: Philosophical Papers* (Cambridge: Cambridge University Press, 1985).

[13] 'What is Human Agency', op. cit., p. 36.

come to a painful end and she has slowly started to enjoy, and depend more and more upon, the company of her old friend, Stolz. Stolz falls in love with Olga and she is confused about her feelings for him.

> If she loved Stolz, then what was her first love? Flirtation, frivolity, or worse? She blushed with shame and turned hot at this thought. She would never accuse herself of that. But if that was her first pure love, what were her relations to Stolz? Again play, deception, subtle calculation, to entice him into marriage so as to cover up the frivolity of her conduct? She turned cold and pale at the very thought of it. But if it was not play, or deception or calculation—so ... was it love again? But such a supposition made her feel utterly at a loss: a second love—eight or seven months after the first! Who would believe her? How could she mention it without causing surprise, perhaps—contempt! She dared not think of it. She had no right. She ransacked her memory: there was nothing there about a second love. She recalled the authoritative opinions of her aunts, old maids, all sorts of clever people, and, finally, writers, 'philosophers of love'—and on all sides she heard the inescapable verdict: 'A woman loves truly only once.' [...] (pp. 400, 401)

Olga concludes that what she feels for Stolz must only be a sisterly love. Stolz confronts her about her baffling behaviour and she is eventually forced to confess that she had been in love with Oblomov and she tells him the whole story of their courtship. When she then shows Stolz a letter Oblomov had written to her very early in their relationship, Stolz uses it to interpret Olga's past feelings rather differently.

> 'Listen,' he said, and he read: ' "Your present *I love you* is not real love, but the love you will feel in the future [...] You have *made a mistake*" (Stolz read, emphasizing the words) "the man before you is not the one you have been expecting and dreaming of. Wait—he will come, and then you will come to your senses and you will feel vexed and ashamed of your mistake" ... You see how true it is,' he said. 'You were vexed and ashamed of—your mistake. There is nothing to add to this. He was right and you did not believe him—that is all your guilt amounts to.' [...]
>
> 'I did not believe him. I thought one's heart could not be mistaken.'
>
> 'Yes, it can, and sometimes very disastrously! But with you it never went as far as the heart,' he added. 'It was imagina-

> tion and vanity on one side, and weakness on the other.' [...] (pp. 411, 412).[14]

When Olga comes to accept this new interpretation of her feelings for Oblomov (albeit mistakenly, I would judge), her love for Stolz becomes a possibility for her in a way that it could not have been without this change of self-conception. She could not rationally hold that she was in love with Oblomov, that a woman only loves once and that she now is in love with Stolz. With Stolz's encouragement she gives up the idea that she had been in love with Oblomov. Once her self-conception or 'practical identity'[15] has been made consistent, she is able to reinterpret her feelings for Stolz as more than mere sisterly love and thereafter allows herself to act freely upon this new conception. The romantic love for Stolz thus takes shape. It is presumably this sort of phenomenon that leads Taylor to claim that

> [w]e can say therefore that our self-interpretations are partly constitutive of our experience. For an altered description of our motivation can be inseparable from a change in this motivation. But to assert this connection is not to put forward a causal hypothesis: it is not to say that we alter our descriptions and then *as a result* our experience of our predicament alters. Rather it is that certain modes of experience of our predicament are not possible without certain self-descriptions.[16]

5. Richard Moran has recently argued, *pace* Taylor, that the sense in which a person's self-conception affects his emotions or other first-

[14] Oblomov wrote the letter that Stolz refers to out of a mixture of cowardice and vanity: partly in an attempt to derail the impending complication that such a relationship would bring to his life, and partly to witness Olga's distress as she reads the letter. His claim that Olga does not really love him is, I think, most implausible, but it is an interpretation with which Stolz can tempt Olga.

[15] The term is taken from Korsgaard, C., *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996).

[16] Taylor, 'What is Human Agency?', op. cit., p. 37.

Compare:

> Our contingent practical identities are, to some extent, given to us—by our cultures, by our societies and their role structures, by the accidents of birth, and by our natural abilities—but it is also clear that we enter into their construction. And this means that the desires and impulses associated with them do not just *arise* in us. When we adopt (or come to wholeheartedly inhabit) a conception of practical identity, we also adopt a way of life and a set of projects, and the new desires which this brings in its wake (Korsgaard, op. cit., p. 239)

order mental states is not a logical, individuative, or constitutive one: it is causal.[17] If his argument is sound, then it might undermine the constructivist view I am attempting to defend and lend plausibility instead to the idea that our mental concepts refer to states that play a complex causal role.

There is a constant slide, Moran argues, between two different stances we might take toward our mental states. On the one hand, we might take what he calls a 'theoretical stance' towards our mental states *qua* independently existing objects and describe or track them. On the other, we might take a practical stance toward them and make a restricted, 'indirect' decision about what to believe or what to intend. This latter, practical question about what to believe is 'transparent to', or answered in the same way as, our theoretical questions about (what is true in) the world.

According to Moran, the rationality of agents has a dual aspect: it ensures that a person's beliefs will aim at the truth, and it ensures that a person's second-order beliefs about his own mental states will affect his first-order beliefs. For example, if the self-interpreter notices an inconsistency or a contradiction in his first-order beliefs his theoretical question about what he believes 'involves reasoning guided by the question of what is true about the object of belief'.[18] The idea, presumably, is that the theoretical question about what I believe will transform itself into a practical question about what *to* believe, since the observation that one's belief is false is, at least *prima facie*, sufficient to destroy the belief. Moran insists that this relationship between second-order and first-order beliefs is not to be construed as a constitutive or logical relationship. It is simply a matter of the tendency of theoretical questions to transform themselves into practical questions. He concludes that 'self-understanding and self-change *can* be understood in a way that maintains the logical independence of interpretation and its object.'[19]

But Moran avoids, I think, the crucial issue in characterizing one of the stances I might take toward my beliefs as *practical* instead of *normative*. On the latter conception, it will be much more difficult to make out a contrast between two different stances, since the normative cuts across the distinction between the theoretical and the practical. There would be some truth, for example, in saying that theoretical questions about what I believe can transform themselves into normative questions about what I should believe, in a way that is transparent to questions about what is true in the world. But

[17] 'Making Up Your Mind', *Ratio*, **1**, (1988), p. 148.
[18] Ibid. p. 148.
[19] Ibid. p. 149.

questions about what I believe also are influenced by questions about what would explain (rationalize) my actions (both past and anticipated present ones), what would best cohere with the other beliefs that I hold or that I have held, as well as my principles, my long term projects, and so forth.

In order to make out his case that beliefs are logically independent of the subject's gaze, Moran suggests that we focus on a person's *past* beliefs and his present interpretations of them, because practical questions do not apply to these and in such cases, the theoretical question about what I believed will not be influenced by practical questions about truth (and thus the question about what *to* believe). The idea is that the theoretical identification of a past belief will not transform itself into a practical question about what *to* believe, since the practical question is now out of date.

But questions about what would explain my actions, what would best cohere with the other beliefs, and so forth, might well influence my identification of a past belief as much as they will influence my identification of a present one. Indeed, because of the pervasiveness of these normative criteria on belief identification, it simply is not clear that there is a viable distinction between a theoretical and a normative stance I might adopt toward my own mental states. (Notice how this point is suggested in the passage from *Oblomov* cited as the epigram to this paper: Olga did not know herself what she thought—what she ought to say or do.) Moran might be right in saying that I can ascribe to myself a past but not a present belief that I know to be false. Nonetheless, because of the other normative constraints (besides the aim for truth) on belief identification, there is plenty of scope to re-introduce the idea that a person's self-conception plays a constitutive, and not merely causal, role in *shaping* (and not merely in describing) his first-order mental states.

In order to pursue this idea further, consider the obligations or entitlements that self-ascription, or indeed any epistemic claim, imposes. A particular ascription, for example, will commit the ascriber to a certain range of justifications he might give for it, if challenged. It will commit him to a range of considerations that would count against it or would follow from it, and included here, of course, would be a certain range of actions. Should a sufficient number of the further commitments fail to obtain or to be endorsed, this puts increased pressure on the self-ascriber to withdraw the original ascription. To take a simple example, suppose that my choice of restaurants might be explained by either the quality of the food or the location. My accepting the latter as a reason, then, ought to affect my attitude toward the suggestion that I might find

equally good food elsewhere. I might, of course, be wrong about why I chose the restaurant (and the alacrity with which I agree to go elsewhere might suggest a reason to suspect that I was). In this case, I ought to re-evaluate my reason for choosing this restaurant. My continued acceptance of location as a reason even if I agree without hesitation to go elsewhere creates a tension which needs sorting either by introducing other reasons for acting into the picture (e.g., I recognize that my companion wants very badly to dine elsewhere) or by my construing my decision to go elsewhere as one that fails to reflect my preference. This is similar to the unstable position Olga found herself in when trying to sort out her feelings for Stolz. My own understanding of or take on my mental attitudes carries with it rational constraints on my future choices, decisions, actions, explanations, criticisms, and justifications. This is true for reasons or attitudes that I attribute to myself as a result of reflection or interpretation; it is also true for immediately ascribed expressions or avowals.

Intentional actions, propositional attitudes, and affective states that presuppose them are identified as such by their role in a pattern of other thoughts and actions. I would like to suggest that if the pattern is a rationalizing one, and thus explanatory in the sense that 'things are made intelligible by being revealed to be, or to approximate being, as they rationally ought to be', then at least many patterns will be 'open-ended' and lend themselves to further, and perhaps different interpretations. Consider how John Wisdom characterizes the idea:[20]

> Suppose two people are speaking of two characters in a story which both have read or of two friends which both have known, and one says 'Really she hated him', and the other says 'She didn't, she loved him'. Then the first may have noticed what the other has not although he knows no incident in the lives of the people they are talking about which the other doesn't know too, and the second speaker may say 'She didn't, she loved him' because he hasn't noticed what the first noticed, although he can remember every incident the first can remember.

Like an aesthetic dispute about, say, the beauty of an object, or a legal dispute about, for instance, whether reasonable care has been exercised, reasons for or against a certain judgment can be adduced. But in cases such as these,

> we notice that the process of argument is not a *chain* of demonstrative reasoning. It is a presenting and representing of those features of the case which *severally co-operate* in favour

[20] 'Gods', *Logic and Language* (Oxford: Basil Blackwell, 1963 (originally published 1951)), pp. 191–192.

> of the conclusion, in favour of saying what the reasoner wishes said, in favour of calling the situation by the name by which he wishes to call it. The reasons are like the legs of a chair not the links of a chain.[21]

This is plausibly the case when we make evaluative judgments in matters of ethics or in practical deliberation about what it would be rational to do as well. Wiggins makes a similar point in a passage in which he is discussing an idea that can be salvaged from a naïve non-cognitivism in ethics and imported into a more sophisticated doctrine of cognitive underdetermination:

> ...not all the claims of all rational concerns or even of all moral concerns (that the world *be* thus or so) need be actually reconcilable. When we judge that this is what we must do now, or that that is what we'd better do, or that our life must now take one direction rather than another direction, we are not fitting truths (or even probabilities) into a pattern where a discrepancy proves that we have mistaken a falsehood for a truth. Often we have to make a practical choice that another rational agent might understand through and through, not fault or even disagree with, but ... make differently himself. ... [22]

It seems to me that the freedom alluded to here with respect to our practical choices figures as well as a feature of our interpretive practices. I suggest that it is arguably indeterminate at the time she broke the lilac whether Olga loved Oblomov just as it was arguably indeterminate whether, at the time she was confronted by Stolz, she had sisterly or romantic feelings for *him*. (I think it is indeterminate whether Constant's Adolphe, in seducing Ellénore, was really in love or was rather simply carried away by the intensity of what had been a game.) Their 'self-takes' play a role in 'articulating' what had been indeterminate before.[23]

[21] Ibid. p. 195.

[22] Wiggins, D. 'Truth, Invention and the Meaning of Life', op. cit., p. 126.

[23] Or, in cases of self-deception, confusing or muddling what had been inchoate before. In these cases there are at least two strands of thought/action patterns manifested. One is the pattern that belies the agent's self-conception and uncovers her ignorance about her own mind. The other is the pattern—often of denials, of protestations, of avoidance—that is a straightforwardly rationalizable outcome of this self-conception. Adolphe, who is self-deceived about the obstacles to his worldy success, is not merely wrong to blame his relationship with Ellénore. His false conception about their life together feeds into a whole pattern of behaviour leading to a tragedy that is itself only rendered comprehensible by this conception.

The case I am making works for states that are inchoate, confused, or multiply interpretable. Not all mental ascriptions or avowals fall into this category. This may be because the subsequent commitments to which an ascription or avowal is answerable have been largely fulfilled, or because it involves relatively little by way of such commitments, like my expression of the desire to have a glass of wine after I have finished work for the day. Other of my desires and intentions—to develop my singing voice, to complete an edging of Bucks Point lace, to expend no more and no less than a reasonable amount of effort doing philosophy—are more complex. They involve commitments extending well into the future that will come into contact and conflict with other intentions, short-term desires, and perhaps some principles or values too. It is true that in considering the application of a concept like love, I am considering complex and pivotal (or central) patterns of action, running from and to numerous other sub-patterns. But although philosophy of mind's discussions tend to take simple, discrete actions (like raising one's arm to signal) as its paradigm case, it is committed to explaining the mental states that figure in an explanation of the projects, plans, commitments, and so forth that constitute a person's life as well.

That it is a *person*'s life is important here. Patterns of animal behaviour can be identified in and rationalized by intentional psychological terms.[24] But although animals can act in accordance with

[24] And perhaps even the 'differential response dispositions' shown by thermometers. Charles Taylor ('What is Human Agency', op. cit., p. 28) suggests parenthetically that Camus's Mersault might an example of someone who fails one test of personhood insofar as he lacks the ability to 'deploy a language of evaluative contrasts ranging over desires' (p. 23). Consider another character from *Oblomov*. Agafya Matveyevna Pshenitzyn, the woman whose elbows entrance Oblomov and eventually capture his heart, is described by the narrator as someone barely capable of self-reflective awareness.

> Had she been asked if she loved him, she would again have smiled and said yes, but she would have given the same reply when Oblomov had lived no more than a week at her house. (p. 374)
>
> ...
>
> He was a gentleman: he dazzled, he scintillated! And, besides, he was so kind; he walked so softly, his movements were so exquisite; if he touched her hand, it was like velvet, and whenever her husband had touched her, it was like a blow! And he looked and talked so gently, with such kindness. ...She did not think all these things, nor was she consciously aware of it all, but if anyone had tried to analyse and explain the impression made on her mind by Oblomov's coming into her life, he would not be able to give any other explanation. (p. 375)

some rational norms (and this might suffice to ascribe intentional states to them) they cannot *follow* those norms. They lack, that is, the meta-ability to understand what the norms commit them to. This will involve an ability to see ways in which a pattern might continue consistently with certain identifications but inconsistently with others. And this ability to recognize patterns and to act in accordance with them because they have been endorsed will introduce a complexity to the patterns that would have been inconceivable for non self-reflective beings. If what I have been arguing here is correct, part of this complexity will involve a kind of self-construction or self-constitution. It is the role of the (whole) person in this construction—the understanding of the commitments and obligations of a rational agent—that seems ill accommodated by causal, reductionist accounts.

6. I have suggested elsewhere that a plausible constructivism about the mental will go some of the way toward explaining the authority of first-person applications of mental concepts and the asymmetry between first- and third-personal ascriptions.[25] A plausible constructivism about the mental will also grant that the choices available in interpretation are not free or unconstrained—anymore than the choices available in musical composition are unconstrained. W. C. Kneale has argued that the important contrast for the constructivist is one that emphasizes the difference between geographical or historical claims on the one hand, and mathematical, and some ethical, practical, and aesthetic claims that are likened to artistic creation on the other.[26] No plausible use of 'invention' in this

[25] 'A Constructivist Picture of Self-Knowledge', op. cit.

[26] 'The Idea of Invention', in *Proceedings of the British Academy* vol. 39, 1955; pp. 85–108.

> ...
> Agafya Matveyevna herself was not only incapable of flirting with Oblomov and revealing to him by some sigh what was going on inside her, but, as has already been said, she was never aware of it or understood it herself ... Mrs Pshenitzyn's feeling, so normal, natural, disinterested, remained a mystery to Oblomov, to the people around her, and to herself. (p.376)

Her brother even characterises her as an animal:

> 'She can't be expected to look after her interests, can she? A cow—that's what she is, a blamed cow: hit her or hug her, she goes on grinning like a horse at a nosebagful of oats.' (p. 357)

context would suppose that it means being capable of creating possibilities from nothing as some—incoherently, he thinks—believe God capable of. 'An artist can do no more than select an interesting possibility'.[27] As long as the distinction between say, geographical and mathematical claims is kept in mind (e.g., that America existed before the first men landed there but the infinitesimal calculus did not exist before it was first formulated), then, Kneale argues, the terms 'invention' and 'finding' may both be apt, since there is no relevant difference between making-with-the-mind and finding-with-the-mind. The contrast between the cases is rather (partly) between what we find with the sense organs and what we find with the mind.

If it is conceded that both the Cartesian and the functionalist suppose that it is something on analogy with a sense organ that 'finds' or 'discovers' the denizens of the mind—the 'mind's eye' on the first model, and an internal scanner on the second—then the contrast is one I can adopt. I am arguing for a rejection of *this* mode of discovery, and am plumping instead for the discovery or selection of something akin to an 'interesting possibility'. This would allow us to begin making sense, then, of the 'compossibility' of objectivity, discovery, *and* invention in the area of psychological discourse.[28]

University of Kent at Canterbury

[27] Ibid. p. 101

[28] This paper was written during study leave made possible by an AHRB Research Award, for which I would like to record my gratitude. It was first presented in March 1999 at a colloquium on Subjectivity hosted by the Kent Institute for Advanced Studies in the Humanities at the University of Kent. Thanks to John Flower, Edward Harcourt and Richard Norman for their comments, and to David Wiggins and Crispin Wright for helpful criticism and suggestions on the penultimate draft.