The Unity of Unconsciousness[1]

Tim Crane

Abstract

What is the relationship between unconscious and conscious intentionality? Contemporary philosophy of mind treats the contents of conscious intentional mental states as the same kind of thing as the contents of unconscious mental states. According to the standard view that beliefs and desires are propositional attitudes, for example, the contents of these states are propositions, whether or not the states are conscious or unconscious. I dispute this way of thinking of conscious and unconscious content, and propose an alternative, which helps to explain why the various mental things that are called unconscious deserve that label.

I *Conscious and unconscious intentionality*

It is widely accepted in philosophy, psychology and cognitive neuroscience that some mental states or episodes or processes are conscious, and many are unconscious. This raises two immediate questions about the relationship between the conscious and the unconscious. First, in what ways do conscious mental states resemble unconscious mental states, and in what ways do they differ? And second, how do consciousness and the unconscious interact? My aim in this paper is to address these two questions.

My interest here is specifically in intentional or representational mental states, states that represent the world. So my question is: how is the way unconscious mental states represent the world related to the way conscious mental states represent it? If we follow the common, if not

universal, habit of calling the way a state represents the world its *content*, then we can rephrase this question thus: what is the relation between the content of unconscious mental states and the content of conscious mental states? The usual answer to this question these days is that these contents are the same kind of thing. The content of an unconscious belief (for example) is a proposition, a bearer of truth-value, and the content of a conscious belief is the same.

We can see that this view might not be totally straightforward, though, when we consider its relationship to the representationalist view that the phenomenal character — that is, the way having an experience seems to a subject in a conscious experience — is identical with its representational content. For, as David Chalmers (2004: §2) has pointed out, how can this be if the content of a conscious belief (say) is the very same thing as the content of an unconscious belief?

Some representationalists will reject the assumption that the content of a conscious intentional state is the same kind of thing as the content of an unconscious state. Tye (1995), for example, argues that the contents of conscious intentional states are abstract, non-conceptual and poised for action, and this is what differentiates them from the contents of unconscious intentional states. One difficulty with this proposal is to explain why a propositional content having these features cannot also be the content of an unconscious state; another difficulty is to explain how states with conceptual content can be phenomenally conscious.

I agree that the contents of conscious and non-conscious states must be different; but I go further than Tye and others, who still hold on to the usual 'propositional attitude' picture for both kinds of state. I will argue here that the way in which unconscious states represent the world is significantly different from the way conscious states represent the world. Concentrating for the purposes of this paper on the state of belief, my proposal is that ascriptions of unconscious beliefs are partial characterisations of what I call a subject's 'worldview' — their entire attitude towards reality. A worldview can be less specific, coherent and determinate than ordinary psychological ascriptions often suggest. These ascriptions serve a modelling and therefore simplifying function, and should not therefore be thought of as corresponding directly to individual belief states with fixed or determinate content.

Where conscious intentionality is concerned, things are very different. Conscious intentional states can also be modelled by propositions, but they have in addition a distinctive way of representing the world which is very different from the way the unconscious represents the world: conscious intentional content is generally more specific, more determinate, more individuated than

the content of the unconscious worldview. But my interest here in this paper is more on the unconscious than the conscious: my central claim is that unconscious states of mind do not have content in the way that conscious states do. This is the first theme of this paper.

My second theme is the nature of the unconscious itself. Many approaches to the mind appeal to the idea of the unconscious — philosophy, cognitive science and psychoanalysis — but there is no clear agreement about what they mean by this. My proposal about the relationship between the conscious and the unconscious provides a story not only about how mental states are brought to consciousness, but also of why all these different approaches to the unconscious are genuinely about the same thing — hence my title.

II Belief: the usual view

I start with the widespread view that beliefs are mental states that represent the world and the way they represent the world is by having representational or intentional content; their intentional content is a proposition, something which is true or false. This view is sometimes summed up by saying that belief states are relations to propositions, and implies that for each belief you have there is a distinct proposition to which you are related.

This picture of beliefs often — though not invariably — goes along with the view that the underlying mechanism for these states is a computational language, a 'language of thought' in Jerry Fodor's famous image. Fodor (1975) put forward his language of thought hypothesis as an empirical claim about the structure of the mind/brain. It is fair to say that in the last forty years psychology and cognitive neuroscience have not discovered anything like a language of thought, though these disciplines use the terminology of mental representation and computation without compunction. Since there are no compelling empirical reasons to believe in a language of thought, I would like to pass over this specific issue and concentrate more generally on the idea of individual belief states.

Beliefs, it is often said, are 'individuated by their contents'. I take this to mean that individual belief states are distinguished from one another by the propositions which give the ways they represent the world. States are standardly thought of as instances of properties or relations, with numerically distinct states being numerically distinct instances. So your and my belief that Princip assassinated the Archduke are usually treated as instances of the same property, or the same relation (the relation expressed by the word 'believes') to the proposition *Princip assassinated the*

Archduke. These are two distinct states, just as this belief of mine is a distinct state from from all my other beliefs. The distinct states are then characterised in various further ways — as dispositions to behave, or to utter sentences, as aiming at truth, as being fine-grained or coarse-grained in their contents, as being externalistically individuated etc. This, in any case, is the textbook picture of belief.

If distinct beliefs are numerically distinct states, then a natural (if naive) question arises: how many beliefs do we have? Given how natural this question is, it is striking how rarely it is asked, and even more striking how few serious answers have been given. Plainly an answer like '1,758,876' seems absurd. When Bob Dylan was asked by a journalist how many protest singers there were, he replied 'either 136 or 142'. His answer was, of course, intended to poke fun at the question. So someone might similarly be inclined to poke fun at the naivety of this question.

Yet I still want to insist that the question should be pursued. For if beliefs really are numerically distinct states, individuated by their contents, why shouldn't there be an answer in principle to the question of how many beliefs you have? What is supposed to be wrong with this question in principle, given the standard view of beliefs? After all, for objects to be individuated is for them to be distinguishable from one another and therefore (in principle) countable. Why shouldn't it be the same with states? And yet there does seem to be something fishy about the idea that you might have a specific, definite number of beliefs, even if it is difficult to figure out what it is. It doesn't seem to be like the question of the number of hairs I have on my head. So rather than dismissing the question and retaining the usual picture of beliefs, perhaps we should consider the obscurity of the question to be a reason for dispensing with the usual picture.

The standard answer to the naive question — in so far as there is one — is that we need to distinguish 'core' beliefs from 'derived' beliefs. Sometimes this is expressed by saying that we have some beliefs which are 'explicit' and some which are 'implicit' or tacit; or some which are 'actual' and some which are 'potential'. These distinctions are not all getting at the same thing, but I will ignore the differences here. For none of these approaches have come up with anything like a criterion, however, vague or imprecise, to distinguish the implicit (etc) from the explicit beliefs. Which are our basic or explicit beliefs? How do we go about settling this? It could be said that this is an empirical question, but even if empirical psychologists were interested in specifying which beliefs are explicit and which are implicit — which to the best of my knowledge they are not — they would still need to start with a clearly implementable idea of what an explicit belief is. And with the

exception of the unworkable idea of a 'sentence in a language of thought', this is what we don't have (see Dennett 1975).

We have our practice of belief-ascription, of course. Some belief-ascriptions are true, others are not. But despite its centrality in our psychological self-conception, our commonsense practice of belief-ascription does not give us any firmer grip on the idea of an explicit belief. An explicit belief is not just one that is consciously available, or one that is easy to ascribe. I can now pluck a belief out of the air, so to speak, and truly ascribe it to myself or to others. I know that I believe that some houses are bigger than others — I only have to consider the question to know that I believe it, and that you believe it too. Does this mean it is a core or explicit or basic belief? Surely not — it's not a particularly important fact for me, it's just clearly a consequence of other things I believe. It's a derived belief if anything is, and yet there it is, immediately accessible for ascription by you or conscious self-ascription by me.

Small children and animals present particular problems in this context, since they cannot put their beliefs into words. As Daniel Dennett says, 'a child can demonstrate his understanding of addition by reeling off sums without being able to formulate or understand propositions about the commutativity of addition. His performance indicates that he has caught on to commutativity, but should we say that among his beliefs is the belief that addition is commutative?' (Dennett 1975: 46). Maybe we should say not, because he does not understand the word 'commutative'. Certainly he does not, but should we then assume that someone has a concept of Fs only when they understand a word for Fs? That isn't part of the standard view.

The view encounters similar difficulties in accounting for the beliefs of animals. We see the dog chasing the cat up the tree, but does the dog's belief involve a proposition about trees in the way that ours does? We think trees are living things, as opposed to houses (say) which are not. So our beliefs about trees are embedded in a network of beliefs that distinguish the living from the non-living, and distinguish them because of (e.g.) the ability of living things to reproduce, our ability to make them grow and so on. Surely the dog has none of this. In what sense, then, does the dog have any beliefs about trees?

But even ignoring children and animals, the standard picture of belief faces questions when it is applied to our effortless (and surely true) everyday ascriptions of beliefs to people. Often we ascribe the same belief to a number of people, using sentences with a common linguistic meaning, even when those to whom the beliefs are ascribed differ significantly in how they take the world to

be. Consider three people, Alice, Bert and Carol who are looking at the Houses of Parliament in London from Westminster Bridge. They all believe that this is the British parliament, but they believe very different things about it. Alice is an expert on the British constitution and has a full knowledge of how the parliamentary system works in the UK. Bert mistakenly thinks that the British parliament is the English parliament, perhaps because he has seen the Scottish Parliament building and because he was brought up in the USA he regularly uses the word 'British' to mean 'English'. Carol thinks that the British government is based in the Houses of Parliament, since she does not distinguish in her mind between government and parliament. These views are so different, the inferences they will draw from their various true and false beliefs are so varied — and yet they all believe that this is the British parliament, as they stand on Westminster Bridge.

How can there be a single belief state that they have in common if the concepts they have — of Parliament, or Britain, or Government — and the consequent inferential properties of their beliefs are so different? Of course there are familiar answers to these questions. We might distinguish, for example, between the core essential (analytic or constitutive) content of the belief and the 'collateral information' associated with its objects; or we might say that these phenomena give us reason to take an externalist or directly referential view of belief content; or we might just give up on the idea that the ascription of the same belief to all these people is literally true. But none of these answers is without its problems.

Some philosophers (most effectively, Stich 1983) have taken these phenomena — animals, children and the holism of belief etc. — to give us reason to dispense with the commonsense notion of belief altogether. But another response is that the problem is not with the notion of belief itself, but with the philosophical conception of it which philosophers have developed from their reflections on the practice of belief ascription. Might there be another way of thinking about beliefs, a way of thinking that better accommodates these facts about belief? To put it another way, if we think that belief ascriptions in these cases are literally true, what is it that makes them true, if not a collection of individual, countable belief states? What is the relationship between the truth of the belief ascription and the psychological reality it describes?

III *Unconscious Belief and the Subject's Worldview*
According to Richard Wollheim, 'the role of belief is to provide the creature with a *picture of the world it inhabits*. Not, of course, any picture of the world, but … a picture that depicts the world

more or less as it is' (Wollheim 1999: 13). The 'picture' of the world which Wollheim talks about is obviously not a conscious picture; we do not have conscious awareness of everything we believe. Philosophers sometimes say that some beliefs are conscious and some are not. In other work I have rejected this idea, and I have argued that all beliefs are unconscious (Crane 2013). There is, of course, the phenomenon of consciously thinking something, but conscious thinking does not amount to believing, even if what one is thinking is what one believes. In the *Philosophical Investigations*, Wittgenstein observed that 'believing is not thinking' and called this 'a grammatical remark', which I take to mean that it is a necessary or conceptual truth (Wittgenstein 1953: §574). My own argument for this conceptual truth was based on the idea that belief is not an occurrence, a point anticipated by Kent Bach thirty-five years ago:

> Philosophers sometimes distinguish between occurrent and dispositional senses of 'believe', but I will use the term 'believe' only for the dispositional sense and reserve the word 'think' for the would-be occurrent sense. I say 'would-be' because I deny that occurrent believing is believing at all, or in my terminology, that thinking that p is either necessary or sufficient for believing that p… Unlike thoughts, beliefs are states, not occurrences. (Bach 1981: 35-4)

What Bach does not say explicitly here is that occurrences of thinking are conscious, and dispositions are not. Clearly, a disposition cannot itself be conscious if it is something that can persist through (e.g.) sleep or other forms of unconsciousness. What can be conscious, though, is the exercise of a disposition — and conscious thinking can be the result of such an exercise. Here I will assume this distinction between belief as unconscious, and thinking as conscious.

The idea that beliefs are dispositions (to behave or to speak) is, of course, an idea familiar from behaviourist and functionalist theories of mind. Roughly speaking, what someone does is determined by what they want, together with the way they believe the world to be. Someone who crosses the room enthusiastically to greet you is doing so because they recognise (and therefore believe) that you are their long-lost friend and want to spend time in your company, or because they recognise you as a long-lost enemy yet want you to think otherwise; or because of some other combination of beliefs and wants or desires. Any one type of action, as Geach (1957) famously showed, could be the product of different beliefs, given different desires.

This gives us the beginnings of an answer to our question about how to reconcile ordinary belief ascriptions with the complexity of individual psychology. I noted above that we are perfectly happy to attribute the belief that this is the British parliament to Alice, Bert and Carol in the situation I described, even though what each of them is disposed to infer from this belief is quite different. As the functionalist story says, the very same belief can give rise to many different other states or actions, depending on the other mental states one has: the belief is a 'multi-track' disposition. In this way, the functionalist story makes sense of the variety of possible outcomes stemming from the same belief state.

But how do propositions — conceived of as abstract objects specifying the contents of beliefs — fit into the functionalist story? My proposal, which I have defended elsewhere (Crane 2015) is that we should think of propositional attitude attributions as models, in the sense described by the philosophy of science (and not in the sense of Tarskian semantics, *pace* Dennett 1982: 154). A model is a (concrete or abstract) object used to draw attention to some features of a system under investigation, and to make the study of the system more tractable. When Rutherford compared the atom to the solar system, this was intended to draw attention to the structural relationship between the atom's nucleus and its electrons. Similarly, a propositional model draws attention to features of your belief, and different kinds of proposition can be used to draw attention to different features. Models idealise and simplify in order to highlight some structural feature of the thing being modelled. Propositions too can simplify — as applied to our example above, we can attribute to different people the belief that this is the British Parliament, when they are standing on the banks of the Thames, even if they have very different (and even incompatible) ideas about what the British Parliament is, or what the Thames is. From this point of view, it doesn't matter that Alice and Bert have such different ideas about Britain. Using the proposition to attribute this belief to them makes it salient that it is *that* building which is the object of their belief, however they happen to conceptualise it.

The case we are considering here is the belief system — the subject's unconscious representation of the world, a representation that aims to 'depict the world more or less as it is', in Wollheim's words. Let's call this the subject's *worldview*. Rather than thinking of each attribution of belief as corresponding to a single belief state, then, let's think instead of the belief system as embodying the subject's whole belief state, their entire doxastic orientation towards the world. The worldview is an unconscious structure that embodies all the dispositions associated with what you

believe. So instead of thinking of a single disposition being associated with a distinct propositional content, the propositions which we employ in belief ascriptions are used to model aspects of this total structure.

Why say that this total structure is something psychological? Some writers (notably John Searle 1992) have argued that there is no unconscious intentionality: there is only conscious intentionality and its neural basis. I reject Searle's way of thinking here, on the grounds that psychology and cognitive science appeal unproblematically to unconscious mental structures and processes (for some examples, see Kihlstrom 1987; Wilson 2002, Hassin, Uleman & Bargh 2005). Treating a creature's worldview as belief amount to this: the creature has a faculty or capacity whose purpose is to represent the world as it is. The worldview has a representational character — it is directed on the world, and its 'direction of fit' is mind-to-world, the terminology Searle (1983) derives from Anscombe. This is a psychological or intentional characterisation of this faculty or capacity, not a neural one, although it is compatible with a further neural characterisation — I leave that question open.

This idea of the worldview is really just another way of expressing the functionalist insight that mental states form a complex network of interrelated dispositional and causal connections — except that I refrain from labelling particular nodes in the network as individual belief states. There is a network of causal and dispositional relations, and we use propositions to model them. This modelling is the essence of individual belief ascriptions, but the truth-maker for these ascriptions need not be individual belief states, as the standard view understands them. I'd like to describe briefly now how this approach gives better accounts of (i) the holistic nature of belief (ii) the beliefs of animals and children, and (iii) contradictory and indeterminate beliefs.

First, the holism of the intentional. This is the fact that a thinker's intentional states must always be part of a wider connected network of intentional states, which influence one another and contribute to their intentional features. We need to hold two facts stable: first, that different people can be ascribed the same beliefs — that is, beliefs with the same propositional contents — and second, that their related states and associated ideas and beliefs may be very different from one another. The worry for the standard of view of propositional attitudes is that if propositions genuinely individuate states, then they must provide identity conditions for the states. This means that any states with the same propositional contents must be identical. But how then can Alice, Bert

and Carol all count as believing that the British Parliament is on the Thames if these beliefs have relational properties — their relations to all their other beliefs —- that are so different?

The picture of propositional attitude attributions as models explains how. The model is only supposed to isolate certain aspects of the subjects' worldviews: that they are representing the UK Houses of Parliament, that they are representing it in a certain way — say, as being on the banks of the River Thames. This is good enough to make the proposition *That is the British Parliament* a model of each of their beliefs, despite the other relational ways in which these beliefs differ. Thinking of propositional attitudes as models explains, then, the phenomenon of unity over difference in our propositional attitude attributions.

Second, the picture gives a better account of what is going on when we attribute propositional attitudes to human infants and animals. Most of us will agree with Normal Malcolm's claim that a dog can believe that a cat went up a certain tree (Malcolm 1973). This can be literally true of the dog. But we should not have to say, because of this, that the dog believes that what the cat went up is a tree in the sense that we believe things are trees. If we treat the proposition, *the cat went up the tree* as a model of the dog's belief, then we can explain its behaviour — hanging around under the tree, barking etc. — by appeal to this proposition. Whatever is the correct account of canine psychology, it is enormously plausible that the dog will have a worldview — its attempt to register the facts about its environment. In anything that counts as a psychological organisation, there must be some component which is responsive to information about the world and whose function is to represent things correctly.

But you may want to insist — 'the cat went up the tree', is that *really* what the dog believes? My answer, assuming my conception of what it is to believe something, is: yes, of course. But this doesn't mean that the content of this part of the dog's worldview cannot be captured in some other way, by using some more theoretical psychological vocabulary. These different ways of describing the same belief should all be compatible, but it is also important that different theoretical purposes will require different levels of determinacy or precision. In their classic discussion of the alarm calls of vervet monkeys, Cheney and Seyfarth claim that 'the meaning of the leopard alarm is, from the monkey's point of view, only as precise as it needs to be' (Cheney and Seyfarth 1990: 169). This might suggest that it is the monkey's decision, as it were, to make its call more or less precise. But of course it is the theorist's decision. A theorist may attribute a 'disjunctive' content to a monkey's

alarm call, but all this means is that this kind of call is provoked by predator X or predator Y; not that the monkey itself has (in some obscure sense) settled on a less committed content.

The third advantage of this approach is that it gives a more realistic account of the obscurity contained in what we actually believe. A lot of what we believe is incomplete, partial, confused and even contradictory. The single proposition-plus-individual belief state picture makes it hard to see how this can be the case, tending to attribute these features to our knowledge of our belief states, rather than to the states themselves. But consider Dennett's poignant example of Sam the art critic, who 'extols, buys, and promotes mediocre paintings by his son' (Dennett 1975: 39). Does Sam really believe that the paintings are good, because he has been blinded by his love for his son? Or does he not believe they are good but pretends they are, out of loyalty and love? Dennett comments that even if a 'neurocryptographer' were to discover an explicit symbolic representation in his brain, this would not necessarily help answer these questions:

> if our neurocryptographer were able to determine that Sam's last judgement on his deathbed was 'My consolation is that I fathered a great artist', we could still hold that the issue between the warring hypotheses was undecided, for this judgement may have been a self-deception. (Dennett 1975: 49)

Let's extend the story a little. Sam is asked to nominate, anonymously, the ten best artists of the up-and-coming generation for a major prize. Let's suppose that the nomination process is so secure and secret that no-one will ever find out whom Sam nominates, and the prize is a significant honour which would advance his son's career significantly. But let's suppose that Sam declines to nominate his son. Is this because he really does not believe that the son is a great artist? Or is it because he does believe it, but he is a man of such integrity that he does not want it on his conscience that he has promoted his son anonymously?

Does Sam, or does he not, believe that his son is a great artist? In a complex case like this, there may be no straightforward yes/no answer to this question. The right description, it seems to me, is that Sam's worldview involves a large collection of dispositions in relation to his son, not all of which are consistent, and the spelling out of which will involve so much complexity that it simply may not be possible to settle on one sentence that describes the belief. Stuart Hampshire has described this kind of situation well:

A man may think that he believes *p*, whilst his behaviour can only be explained by the
hypothesis that he believes not-*p*… Perhaps the confusion in his mind cannot be conveyed
by any simple account of what he believes; perhaps only a reproduction of the complexity
and confusion will be accurate. (Hampshire 1975: 123)

So we need to be able to say that it may simply be indeterminate whether Sam believes that his
son is a great artist. But this is not because there are no psychological facts about what he
believes — it's rather because there are too many. Complexity and confusion can go right to the
bottom of our worldview.

Not all cases are like this, of course. Many aspects of our worldview are straightforward and
the dispositions involved can be modelled effectively with a single proposition. But even in these
straightforward cases, the holism of the intentional guarantees that there need not be one single
dispositional profile associated with each proposition believed.

Modelling by a third party is one way that a determinate content can be imposed upon the
worldview. The other is, of course, self-interpretation: introspection or the self-attribution of beliefs.
In the recent philosophical literature, this subject is often discussed under the heading of 'self-
knowledge'. As Quassim Cassam points out (2015: vii), this is not an entirely happy label, since the
knowledge discussed is not normally knowledge of one's *self,* but knowledge of one's beliefs and
other attitudes. I agree; but here I want to focus on another aspect of self-knowledge in this sense.

Often when we are asked what we believe, we can answer the question straight off — 'do
you believe that it is raining now?' and so on. But there are some cases where we have to work out
what we believe. Often when we reflect on our beliefs about some complex matter — a political or
theoretical question, for example, or a situation involving our relationships with those close to us —
we find ourselves with confused and often contradictory or unresolved attitudes. (The situation is
the same for what we want or desire, but I am focussing only on belief in this paper.) The picture of
'self-knowledge' suggests that there are determinate facts about our beliefs and we just need to
employ some method to find out what they are. But often when we introspect (whatever that
exactly means) we find only confused and unclear beliefs. In these kinds of case, bringing our
beliefs to consciousness may require us to resolve this confusion, by deciding to express ourselves
in one way or the other. In this way, the conscious acknowledgement or expression of a belief has

a different kind of content — more fixed, more determinate — than the content of much of the worldview.

Again, it's easier to see what I mean here by contrasting it with the picture of fully determinate belief states (whether understood as 'sentences in the head' or as fully determinate dispositions). That picture implies that all the facts about our beliefs are there waiting to be discovered, and that the difficulty we have in articulating them is purely epistemological. On this view, discovering what you think is one thing; making up your mind is entirely another. The alternative I am proposing is that there is no sharp distinction between these two activities. Discovering what you believe can resolve indeterminacy and unclarity in your worldview, and producing a conscious judgement that settles things as far as you are concerned. This view treats the determinate content of a belief as 'coming into being as we probe', as Michael Dummett once said of a different kind of case (mathematical reality: Dummett 1959).

The clearest way in which the content of a worldview can become the content of a conscious judgement is when one says something, either aloud or to oneself. This is the truth behind the often-quoted aperçu, 'How can I tell what I think till I see what I say?', attributed variously to E.M. Forster (1927: 99) and Graham Wallas (1926). When we use words to express our beliefs, often by uttering a sentence, we impose an order on what we think which can then make our subsequent dispositions more determinate. We need to recognise how, in conscious thinking, 'the hard edge of determinacy our verbal output substitutes for the fuzziness of our convictions' as Dennett (1975: 48) nicely put it. But we can also do this without putting it into words — we can make a conscious judgement without explicitly asserting a sentence.

The central phenomenon here is interpretation. To model an aspect of a worldview with a proposition is to interpret it. To bring a belief to consciousness is to interpret something about your worldview, where interpretation is a matter of taking a system — you — and making sense of how you represent the world in general. From the third-person point of view, you take the subject's belief system or worldview and isolate some feature of it by relating it to a proposition. From the first-person point of view, recovering what your beliefs are is a matter of interpreting yourself — finding out what is the best way to express or describe your beliefs. This can involve making up your mind, or figuring out what is straightforwardly contained within it anyway. Interpretation can bring determinate content to consciousness, or it can render determinate what is not yet determinate.

Saying this is not the same as endorsing an 'interpretationist' conception of intentionality in general — as if all facts about intentionality can be explained by appealing to the idea of interpretation. (William Child (1996) attributes this view to Davidson and Dennett.) For one thing, interpretation itself is a mental or intentional activity, so we cannot expect all intentionality to be explained in terms of interpretation. The interpretation of one's own belief is a state or episode of consciousness, and this has its intentionality independent of any further act of interpretation. Conscious acts of interpretation do not require interpretation to give them content.[2]

IV *The Unconscious in General*

How do these ideas about belief relate to the unconscious in general? The notion of unconscious mentality figures largely in at least three distinct theoretical areas: cognitive science, the philosophical interpretation of commonsense psychology, and psychoanalysis. It is worth asking whether the notion is used in the same way in all these areas. Is there a common, unified notion of the unconscious?

With a few notable exceptions, psychoanalysis has not received the kind of discussion in analytic philosophy of mind as cognitive science and commonsense psychology have.[3] So it is an interesting fact that in their introductory expositions of the unconscious, both philosophers and cognitive psychologists frequently credit Freud with the insight that there are unconscious intentional states. Here are a few examples. Jerry Fodor says:

> It used to be universally taken for granted that the problem about consciousness and the
> problem about intentionality are intrinsically linked: that thought is *ipso facto* conscious, and
> that consciousness is *ipso facto* consciousness of some or other intentional object… Freud
> changed all that. He made it seem plausible that explaining behaviour might require the
> postulation of intentional but unconscious states. Over the last century, and most especially

---

[2] This is one respect in which the views of this paper depart from those of Dennett, to whom I am otherwise deeply indebted. In *Consciousness Explained* (1991), for example, Dennett defends the view that even the content of states of consciousness is determined by a certain kind of interpretation.

[3] Some of the notable exceptions are Hopkins and Wollheim (1982) and Gardner (1993). The situation is somewhat different in so-called continental philosophy, perhaps because of the extensive influence of Jaques Lacan.

in Chomskian linguistics and in cognitive psychology, Freud's idea appears to have been

amply vindicated. (Fodor 1991: 12)

Much more recently Peter Carruthers has observed that 'almost everyone now accepts … (post-Freud) that beliefs and desires can be activated unconsciously' (2016: §1). In psychology too, Freud often gets the credit for the notion of the unconscious: 'the idea that a large portion of the human mind is unconscious' writes Timothy Wilson 'is not new and was Freud's greatest insight' (Wilson 2002: 3).

And yet Freud's specific ideas themselves are largely ignored in both analytic philosophy of mind and in psychology.[4] After acknowledging Freud's insight, Wilson goes on to say that 'the modern adaptive unconscious is not the same as a the psychoanalytic one' (Wilson 2002: 5). Psychoanalysis as a theoretical enterprise operates largely outside the mainstream of psychology and philosophy — as I see it, there is more communication between cognitive science and mainstream philosophy of mind than there is between either endeavour and psychoanalysis. But this raises a puzzle — if psychologists and philosophers of mind ignore Freud's ideas, why do they pay tribute to him for his great discovery? What are they actually crediting Freud with? Or to put it another way: what did Freud actually discover?

Not the unconscious, as is now widely acknowledged. If any one thinker should be credited with making explicit the idea of unconscious intentionality, it is Leibniz, who in his *New Essays* (1704) talked about the *petites perceptions* which are 'either too minute and too numerous, or else too unvarying' to attract our awareness (Leibniz 1996: 53). Leibniz drew attention to these phenomena in part to undermine the Cartesian view that the soul is always conscious. Also influential was Hermann von Helmholtz, who argued that perception must be explained in terms of unconscious inference — disciplined transitions among intentional states (Helmholtz 1853). Other 19th century thinkers like Gustav Fechner (1801-87) and Friedrich Albert Lange (1828-75) made use of ideas of unconscious mentality in their theorising. Freud was not the first.

---

[4] Katalin Farkas (forthcoming) provides some evidence on the philosophy side: 'at the time of writing [October 2014], there are around 1500 entries in the *Stanford Encyclopedia of Philosophy*…. 92 of these entries (mostly on continental or feminist philosophy) refer to Freud, whereas for example Hilary Putnam is mentioned in 230 entries. The *SEP* entry on the "Philosophy of Psychiatry" does not contain a single reference to Freud's work. The term "psychoanalysis" is mentioned in 76 documents. In contrast, "cognitive science" occurs in 175 documents, "artificial intelligence" in 119 documents, ""quantum mechanics" in 138 documents.'

Paul Katsafanas has argued recently that a distinction can be made between those theories of the unconscious which treat it as belonging to essentially of the same kind as the conscious, and those which see it as something of a very different kind — even as something opposed to, or set against, consciousness (Katsafanas 2016: 17-21). As Katsafanas says, some theories of the first kind treat consciousness itself as something which is the upshot of the combination of unconscious states. Leibniz, for example, argued that when some *petites perceptions* are 'combined with others they do nonetheless have their effect and make themselves felt, at least confusedly, within the whole' (Leibniz 1996: 53). But those that treat the conscious as essentially the same kind of thing as the unconscious need not think of the former as constituted out of the latter; there are other ways to think of the unconscious on this model.

Freud's view of the unconscious is of the second kind distinguished by Katsafanas: the unconscious is 'a system with its own rules and principles' (Wollheim 1971: 13). One central Freudian theme, for example, is that of repression. Ideas which cannot be admitted into consciousness are repressed; these ideas are resisted by the conscious mind, with resistance being lowered only in the context of dreams and in therapy. There is a struggle between the unconscious drives and the conscious resistance of their products; and what is moving things in the unconscious is very different from what moves us consciously. This is a very different conception of the unconscious from that which treats unconscious intentional states as essentially the same kind of thing as conscious states, just not as powerful (as Leibniz did). Looked at in this way, Freud's original achievement was not simply to draw attention to the distinction between intentionality and consciousness — as Fodor, Carruthers and many others claim — but rather to defend a conception of the unconscious as something of a wholly different nature from consciousness.

The contemporary conception of the relationship between conscious and unconscious intentionality is, in this respect, more similar to Leibniz's than to Freud's. Unconscious propositional attitudes represent the world in the same way that conscious ones do — by having propositional content, that is their essence as intentional states. What makes them conscious is something else — qualia, higher-order representation etc. — which is not anything to do with their intentional content (how they represent the world). In rejecting here this conception of the relationship between conscious and unconscious intentionality, I am therefore thinking of the unconscious in broadly Freudian rather than Leibnizian terms.

My aim here, however, is not to defend Freud's specific view of the mechanisms of the mind, with its postulated structures of id, ego and superego. I do think that many of the phenomena psychoanalysis describes are real — at least in the sense that among the deepest human motivations are certain drives that they do not fully understand or recognise from simple or superficial introspection, which often result from past experience now forgotten. The aim of psychoanalytic therapy is to uncover these motivations, and whether this kind of therapy is generally effective is a question I cannot consider here. What I want to draw attention to instead is how the psychoanalytic conception of the relationship between the conscious and the unconscious — whatever its merits as a theory or as therapy — conforms to the picture I sketched in the previous section.

The key point in common, again, is interpretation. The psychoanalytic process (or indeed, any therapy aimed at uncovering drives or desires which are not immediately manifest) involves the patient reflecting on things that have happened to them, or things that they have done, in order to expose the repressed desires or drives, and make their behaviour more intelligible to them. This essentially involves creating a narrative of their lives and revealing those ideas or drives which are repressed and play a role in the explanation of their behaviour. I agree with James Hopkins, then, when he argues that psychoanalytic interpretation is, in a certain way, a process which is continuous with that of commonsense psychology (Hopkins 1981). But this is not because psychoanalytic interpretation simply takes on board the standard philosophical conception of the propositional attitudes and their contents, and applies it to the psychoanalytic unconscious. It's rather because commonsense psychological interpretation essentially involves the kind of mixture of discovery and imposition of structure which is also there in the psychoanalytic session. Just as 'self-knowledge' can involve making up your mind about what you believe, so the recovery of unconscious drives can be part of a story that you tell to make you intelligible to yourself. Your unconscious drives are part of your unconscious motivational system. Your attempts to introspect your worldview can involve creation as well as discovery, and the same applies to your attempts to recover your motivations by interpretation. (Remember that I am restricting the term 'worldview' to your belief system; the motivational system is something distinct, although many of the same things are true of it.)

My hypothesis is that cognitive science, commonsense psychology and psychoanalysis all operate with a conception of the unconscious as something that has a causal structure which is

modelled by the attributions of propositional content by the theorist or the therapist. In order for these attributions to provide a successful interpretation of the unconscious mind there is no requirement that they correspond in any straightforward sense to individual inner states with determinate propositional contents. The same is true of the subject's own first-person reflections on their worldview, and on their unconscious motivations.

V *Conclusion*

The standard philosophical picture of the propositional attitudes is a mixture of important truths — the functionalist truisms about dispositional profiles, the ideas of direction of fit and representational content — and some unrealistic and misguided dogmas, such as real relations to abstract propositions, the pursuit of a compositional semantics for intentional states, or for sentences in a language of thought. The latter ideas have recently have recently come under attack because of their inability to handle consciousness — for surely any plausible story about consciousness must incorporate a story about the subject's intentional representation of the world? The way to approach this question, I have argued here, is to abandon certain aspects of the standard picture of the propositional attitudes, and to work out an alternative picture of unconscious intentional content. It should be a constraint on any such picture that it makes sense of the way the unconscious is conceived across many different kinds of theory of the mind.

Faculty of Philosophy

Sidgwick Avenue

Cambridge

CB3 9DA

tc102@cam.ac.uk

www.timcrane.com

References

Bach, Kent (1981) 'An Analysis of Self-Deception' *Philosophy and Phenomenological Research* 41: 351-70.

Carruthers, Peter (2016) 'Higher-Order Theories of Consciousness', *The Stanford Encyclopedia of Philosophy* (Fall 2016 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2016/entries/consciousness-higher/>

Cassam, Quassim (2015) *Self-Knowledge for Humans* (Oxford: Oxford University Press).

Chalmers, David (2004), 'The Representational Character of Experience' in Brian Leiter (ed.) *The Future for Philosophy* (Oxford: Oxford University Press).

D.L. Cheney and R.M. Seyfarth, *How Monkeys See the World*: *Inside the Mind of Another Species* (Chicago, Ill.: University of Chicago Press 1990).

Child, William (1996) *Causality, Interpretation and the Mind* (Oxford: Oxford University Press).

Crane, Tim (2013), 'Unconscious Belief and Conscious Thought' in *Phenomenal Intentionality: New Essays* edited by Uriah Kriegel (Oxford: Oxford University Press 2013) 156-73.

Crane, Tim (2015) 'The Mental States of Persons and their Brains' in Anthony O'Hear (ed.) *Mind, Self and Person, Royal Institute of Philosophy Supplement 76* (Cambridge: Cambridge University Press 2015) 253 - 270.

Dennett, Daniel C. (1975), 'Brain Writing and Mind Reading' in Keith Gunderson (ed.) *Minnesota Studies in the Philosophy of Science* 7:403-15, reprinted in Dennett, *Brainstorms* (Hassocks: Harvester 1981).

Dennett, Daniel C. (1982), 'Beyond Belief' in Andrew Woodfield (ed.) Thought and Object (Oxford: Oxford University Press).

Dummett, Michael (1959) 'Truth' *Proceedings of the Aristotelian Society* 59:141-62.

Farkas, Katalin (forthcoming) 'The Boundaries of the Mind' in Amy Kind (ed.) *Routledge History of the Philosophy of Mind: the 20th Century* (London: Routledge).

Fodor, Jerry A (1975) The Language of Thought (Hassocks: Harvester).

Fodor Jerry A (1991) 'Too Hard for our Kind of Mind' *London Review of Books* 13, June 27.

Forster, E.M. (1927) *Aspects of the Novel* (reprint Harmondsworth: Penguin 1976).

Gardner, Sebastian (1993) *Irrationality and the Philosophy of Psychoanalysis* (Cambridge: Cambridge University Press).

Geach, P.T. (1957) *Mental Acts* (New York: Humanities Press).

Hampshire, Stuart (1975) *Freedom of the Individual* (Princeton, NJ: Princeton University Press).

Hassin, Ran, and James S. Uleman & John A. Bargh (eds.) (2005) *The New Unconscious* (Oxford: Oxford University Press).

Hopkins, James (1982) 'Introduction' in James Hopkins and Richard Wollheim (eds.) *Philosophical Essays on Freud* (Cambridge: Cambridge University Press).

Katsafanas, Paul (2016) *The Nietzschean Self* (Oxford: Oxford University Press).

Kihlstrom, John F. (1987) 'The Cognitive Unconscious' *Science* 237:1445-1452.

Leibniz, G.W. (1996) *New Essays on Human Understanding* (Oxford: Oxford University Press).

Malcolm, Norman (1973) 'Thoughtless Brutes' *Proceedings and Addresses of the American Philosophical Association* 46: 5-20.

Searle, John (1983) *Intentionality* (Cambridge: Cambridge University Press).

Searle, John (1992) *The Rediscovery of the Mind* (Cambridge MA; MIT Press).

Speaks, Jeffrey (2015) *The Phenomenal and the Intentional* (Oxford: Oxford University Press).

Stich, Stephen P. (1983) *From Folk Psychology to Cognitive Science* (Cambridge, MA: MIT Press).

Tye, Michael (1995), *Ten Problems of Consciousness* (Cambridge, MA: MIT Press).

Wallas Graham *The Art of Thought* (London: Jonathan Cape 1926).

Wilson, Timothy (2002) *Strangers to Ourselves* (Cambridge, MA: Harvard University Press).

Wittgenstein, Ludwig (1953) *Philosophical Invesitgations* (Oxford: Blackwell).

Wollheim, Richard (1971) *Freud* (London: Fontana).

Wollheim, Richard (1999) *On the Emotions* (New Haven: Yale University Press).