SS4 Readings on Consciousness

2. Jackson, F. (1982) 'Epiphenomenal qualia', *Philosophical Quarterly*, 32.
JSTOR via TCD local library page.

3. Levine, J. (1993), 'On leaving out what it's like', in *Consciousness*, M. Davies & G. Humphreys, eds, Oxford: Basil Blackwells, reprinted in *The Nature of Consciousness*, N. Block, O. Flanagan & G. Güldezere eds, Cambridge MA: MIT Press.

4. Rosenthal, D. (1990), 'A theory of consciousness', Report 40, 1990, Research Group on Mind and Brain, Center for Interdisciplinary Research (ZiF) at University of Bielefeld, Germany, reprinted in *The Nature of Consciousness*.
Web site:http://web.gc.cuny.edu/cogsci/th.htm

5. Lycan, W. (1990) 'Consciousness as internal monitoring' in J. Tomberlin (ed.), *Philosophical Perspectives*, vol.4; repr. in a revised form in *The Nature of Consciousness*.
Web site: http://www.nd.edu/~lstubenb/Lycan.html

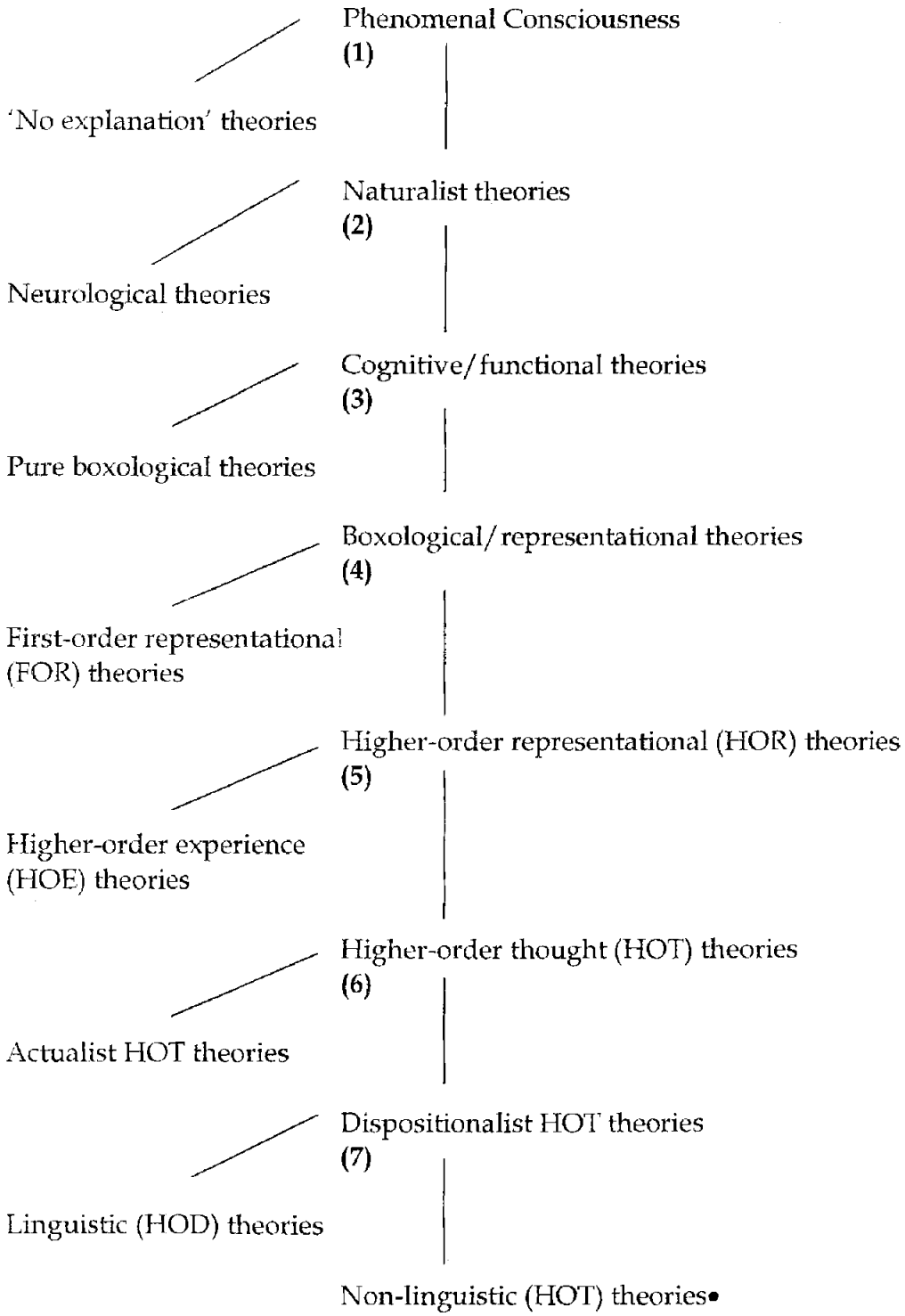
6. Dretske, F. (1993), 'Conscious experience', *Mind*, 102: 263-283, reprinted in *The Nature of Consciousness*, N. Block, O. Flanagan & G. Güldezere eds, Cambridge MA: MIT press.
Web site: JSTOR via TCD local library page.

7. Block, N (1995) 'On a confusion about a function of consciousness', *Brain and Behavioural Sciences*, 18, reprinted in *The Nature of Consciousness*, N. Block, O. Flanagan & G. Güldezere eds, Cambridge MA: MIT press.
Web site: http://www.nyu.edu/gsas/dept/philo/faculty/block/

8. Carruthers, P. (1998), 'Natural theories of consciousness', *The European Journal of Philosophy*, 6.
Website:http://www.philosophy.umd.edu/people/faculty/pcarruthers/Natural-consciousness.htm

Also available via philosophy undergraduate courses page. Full bibliography can be found at David Chalmers website:
http://www.u.arizona.edu/~chalmers/online1.html#higher

Choice Points for a Theory of Consciousness



SS4 Consciousness

Seminar 2: The Explanatory Gap

The problem of the explanatory gap stems directly from the knowledge argument: if one does not know all there is to know in knowing all the physical facts, i.e. there are additional phenomenal facts, then there is an explanatory gap between the physical and phenomenal facts. It also stems from Nagelian considerations: we do not have the first understanding of how to relate facts about conscious experience with objective facts.

Levine first discusses what he takes to be the real import of the intuition behind both the modal argument and the knowledge argument. Kripke and Jackson alike intend their arguments to have metaphysical implications, but, according to Levine, they only have epistemological implications. Nevertheless Levine argues that there is something damaging to physicalism about these two arguments, viz. they highlight a weakness in the explanatory capabilities of physicalism.

According to Kripke neither standard scientific identity statements nor mental-physical identity statements are known *a priori*; therefore they are both imaginably false, that is, they seem to be contingent. Yet, if they are true, they are necessarily true, i.e. they are not even possibly false. So how do we reconcile the apparent contingency with the actual necessity? Following Kripke this is clear in the first case; we think, for example, we are imagining a situation in which water is not H_2O, but in fact we are imagining a situation in which some substance which behaves superficially like water, but is not water, is not H_2O. (Note: one can always redescribe the original situation in such a way that one is giving a true description of something else.) But a similar explanation cannot be given for the apparent contingency of the identity between a mental state and a brain state, because to imagine a situation in which one is experiencing a state superficially like pain just is to imagine a situation in which one is experiencing pain.

But a metaphysical distinction does not automatically follow; perhaps mental and physical concepts do refer to the same property. (Typically identity statements are informative because we track two properties to the one object; the physicalist will have to say that there are not two properties but two ways of speaking about physical properties.) But epistemological consequences do seem to follow.

Levine sees the disanalogy between the two kinds of cases in the way that whereas there is an apparent necessity that flows from the reduction of water to H_2O, this is missing from the reduction of, for instance, pain to the firing of C-fibres. The reason for this is that whereas the apparent contingency of the first kind of case only works in one direction (while it is conceivable that something other than H_2O should manifest the superficial properties of water it is not conceivable that H_2O should fail to manifest these properties) it works in both direction in the second kind of case (it is equally conceivable that there should exist a pain without the firing of C-fibres, and the firing of C-fibres without pain). Whereas we have a chemical theory that explains the macro-properties of water, we have no such theory for the equivalent properties of mental states, such as the painfulness of pain; an explanatory gap exists between physical states and mental states.

Levine's conclusion is that concepts referring to conscious experiences and their properties are - as yet - inappropriate for relating to underlying functional physical properties.

Conceivability Argument and the Explanatory Gap Argument

Philosophers now tend to distinguish these two arguments. Conceivability arguments focus on concepts (and ontology); explanatory gap arguments focus on laws (and epistemology). One cannot relate phenomenal and physical concepts *a priori* because they are different concepts. Hence we can imagine what they refer to as distinct properties. But it may be that what they refer to can be related empirically and that they are identical. Chalmers has provided an influential discussion using supervenience. He claims that we can imagine the mental existing independently of the physical because the former supervenes naturally, as opposed to logically, on the physical. Biological and chemical properties supervene logically on the physical. He motivates his view by epistemic considerations (hence the link to explanatory gap arguments). We can know the logically supervening properties if we are given the logically subvenient properties, but we cannot infer the merely naturally supervenient properties from the naturally subvenient properties. Here are two ways to respond to conceivability arguments:

A. Argue that we cannot in fact ever infer supervenient properties from subvenient properties (Levine's recent view).

B. Argue that phenomenal concepts are special. Unlike physical properties which are functional, phenomenal concepts pick out non-functional properties.

Pessimistic Views of the Explanatory Gap

McGinn believes that there is a natural explanation for consciousness but we are cognitively closed to it. Suppose that property P is what gives rise to consciousness and theory T is an understanding of P such that would explain the mind-body problem. There are two ways we might understanding T: via introspection on our experiences or via perception of the brain. But neither of these are adequate to the task. Introspection only gives us access to our mental lives, and although we can infer things beyond our immediate perception, which then explain those perceptions, these will be limited to certain kinds of explanations, specifically in terms of spatial concepts. Consciousness is not spatial. Hence we are cognitively closed to an explanation of consciousness.

Optimistic Views of the Explanatory Gap

The explanatory gap is merely a temporary lacuna in our understanding. But we should not expect to find a direct reduction of the mental to the physical. Rather we should try to work out what consciousness involves in terms of functional/computational processes and then relate these to physical properties.

The Explanatory Gap as a Cognitive Illusion

Tye holds that the explanatory gap is a kind of cognitive illusion brought about by the failure to recognize the special features of phenomenal concepts. Phenomenal concepts (the concepts utilized in introspecting on phenomenal states) are irreducible concepts: each phenomenal concept is tied to a particular experience-specific perspective occupied by the possessor of the concept.

The Explanatory Gap and the Antipathetic Fallacy

A related response is provided by Papineau. We continue to think that the phenomenal and the physical are distinct, despite their identity, because of the different function of phenomenal and physical concepts; the use of phenomenal concepts generate related experiences, the use of physical concepts does not, and we mistakenly infer from this that there is a distinction in the properties themselves.

SS4 Consciousness

Seminar 4: Higher Order Thought (HOT) Theory of Consciousness

Philosophy of mind is often divided into two: the problem of the sensory and the problem of the intentional. Rosenthal suggests a further division: a study of consciousness independently of the sensory and the intentional.

Rosenthal starts by making an important and influential set of distinctions. Any account of consciousness must be cognisant of the different uses of the term 'consciousness'. Firstly, one might talk about *creature consciousness*. This notion can be understood relatively unproblematically as the difference displayed by creatures when awake rather than asleep. Creature consciousness involves the creature being conscious of things around it. Secondly, one might talk about *state consciousness*. This notion concerns mental states, especially perceptual states. It is this form of consciousness which philosophers now recognize as requiring explanation. Rosenthal distinguishes a third notion of consciousness: *introspective consciousness*. This, he holds, does not generate any special problems different from those already associated with non-introspective consciousness. It is the structure of state consciousness that has to be understood before we can hope to make any progress with the problem of the explanatory gap between the physical and the phenomenal. Our lack of progress has made us attribute all sorts of properties to consciousness, unanalyzability, transparency, immediacy, epistemic privilege etc.

Rosenthal claims, against Descartes, that not all mental states are conscious. We often do think in terms of unconscious states, both intentional and sensory. But there also seem to be explanations for the Cartesian intuition. We normally do focus on our own and others' conscious states (more so with sensory states than intentional states). But our normal lack of interest in non-conscious states should not seen as a guide to their lack of existence. We might try to explain non-conscious states as mere dispositional states does not meant that they are not also occurrent states. We may originally distinguish mentality by reference to the phenomenal character (the what it is like) of conscious states, but this does not imply that all mental states, nor even the archetypical conscious states must be conscious; the distinctive qualities by which we type sensations could be possessed by non-conscious sensations. Consciousness may be a single mark that is used to fix the reference of the mental, but it may not be an essential property of the mental, and there are some reasons to think that the intentional or the sensory are more fundamental.

Rosenthal claims further that an answer to the problem of consciousness will only be forthcoming if we assume that some mental states are unconscious. Firstly, if all mental states are conscious we cannot have a non-circular explanation of conscious mental states in terms of other mental states. He claims that we will be better able to relate the mental to the physical if we first consider the relationship between the conscious and non-conscious. Secondly, that consciousness is intrinsic (necessary) to mental, especially sensory, states is appealing, but we would have no hope of understanding consciousness if it is an intrinsic rather than a relational property of

mental states. Rosenthal specifically emphasizes that sensational properties (another way of thinking of qualia) can exist unconsciously. This is an idea which may be hard for many, who like Kripke claim that the feeling of painfulness is an essential property of pain, to accept.

Another distinction is first required in order to explain how some states are conscious and other states are not: between *transitive consciousness* and *intransitive consciousness*. The former involves being conscious of something. The latter does not. Conscious states, for Rosenthal, only have intransitive consciousness. Rosenthal's theory states that what makes some mental states intransitively conscious and others not is a specific kind of relation between mental states. More specifically a first-order mental state is intransitively conscious if another second-order transitive mental state is directed on to it, i.e. we are transitively conscious of former with the latter. The higher-order state cannot be intrinsically conscious; it is this which saves Rosenthal's theory from circularity. And transitively conscious states can occur without intransitive states. Transitive consciousness is the consciousness that creatures have. In Rosenthal's account a necessary condition of the higher-order state is that it be non-inferentially related to the lower-order state (though not necessarily directly caused by the lower-order state).

The higher-order states might be perceptions or thoughts. The first is favoured by comparison with perceptual states: sense modalities respond differentially to a range of stimuli as being conscious might respond to a range of mental states; sense modalities respond directly and consciousness is supposed to be immediate; and both seem similarly inexplicable. Conversely, mental states do not exemplify a single range of stimuli as sensory modalities do; and they do not have a specific character of experience as perceptual experiences do. Thinking, by contrast, is not so limited to a single range of stimuli; it is just as unmediated; and it does not have characteristic qualities. This is the higher-order thought (HOT) theory of consciousness. HOTs are constrained; we are conscious of our mental states in a special way: as our possessing them. HOTs do not require complex conceptual abilities. Animals might have simple HOTs. Finer-grained distinctions may be provided by the conceptual resources of language. Dispositional states are insufficient. But this theory seems to require a multitude of HOTs to account for all our conscious states. As against this objection we clearly do have many non-conscious thoughts and we may not need that many HOTs anyway. So, some mental states are intransitively conscious because of the way the thoughts we have about them make us transitively conscious of them.

Introspection is the directedness of a third-order state on to a second-order state, making the latter intransitively conscious. Hence introspection is explained in terms of a prior account of state consciousness, and further motivates the distinction between transitively and intransitively conscious states.

SS4 Consciousness

Seminar 7: Representational Theories of Consciousness

Whereas a higher-order theory of consciousness holds that a mental state is conscious in virtue of another mental state (or monitor) being directed on to it, a first-order representationalist theory of consciousness holds that mental states can be conscious states in their own right. But they are so in virtue of representing external things (i.e. contents). The two main exponents of such a view are Fred Dretske (*Naturalizing the Mind*) and Michael Tye (*Ten Problems of* Consciousness).

Representationalism: Phenomenal Character and Consciousness

The main representationalist thesis about the character of experience (what an experience seems like to its subject) is that it is identical to the content represented by an experience. Representationalists are typically 'externalist': the content represented by an experience that constitutes the character of an experience is outside the head. (Ned Block has claimed that 'the greatest chasm in the philosophy of mind - maybe even all of philosophy - divides two perspectives on consciousness [...] whether there is anything in the phenomenal character of experience that goes beyond the intentional, the cognitive and the functional'.)

Tye holds that necessary and sufficient conditions for consciousness are given by his PANIC theory. States are conscious iff they are constituted by Poised Abstract Nonconceptual Intentional Content. The content is then individuated by causal relations with the external world:

> A state, S, represents that P =def. if optimal conditions obtain, S is tokened in [subject] x if and only if P and because P.

Dretske holds that states are conscious if they have the function (are designed) to represent certain properties:

> A system, S, represents a property, F, if and only if S has the function of indicating (providing information about) the F of a certain domain objects.

Dretske emphasizes that such systems, having indicator functions, are natural (rather than conventional), and non-conceptual, and can be used by the subject in higher-cognitive activity.

According to both philosophers (and most representationalists) if a creature is in a state that represents in roughly the way outlined above, and this representational state is available for further cognitive processing, then the state will be a conscious one. On such a view cognitive processes can be broadly divided into two main kinds: modular processes which generate representational contents in a mechanical fashion, and processes of the belief/desire system which operate on the contents provided by modular processes.

Conscious Experiences

Dretske starts by drawing our attention to another important distinction, that between *thing awareness* and *fact awareness*. We can be thing aware of something without being fact aware that something is the case and vice versa; the difference lies in whether the mental state is charged by a concept. For instance, we can be aware of Clyde playing the piano (e.g. by hearing him) without being aware that Clyde is playing the piano (e.g. we do not know that it is Clyde that is playing the piano), or conversely we can be aware that Clyde is playing the piano (e.g. someone has told us) without being aware of Clyde playing the piano (i.e. we cannot hear him ourselves).

Dretske's argument advances by steps. To see something it is natural to think that one is conscious of that thing or, if one has the appropriate facts at one's disposal, to see that something is the case is to be conscious of the fact (a conscious belief):

1. S sees (hears etc.) x (or that P) => S is conscious of x (or that P).

But it does not follow from being conscious of something that one is conscious of a fact about that thing; one needs the appropriate concepts to be conscious in that way. A momentary reflection on the way our grasp of new concepts makes us newly aware of facts should help to make this clear. So:

2. S is conscious of x $\not\Rightarrow$ S is conscious that x is f.

So one can be conscious of the thing that is the difference (thing aware) without being conscious of the difference (the fact that there is a difference).
Turning to consciousness itself. we have the commonsense view that transitive consciousness implies intransitive consciousness (though not necessarily its converse):

3. S is conscious of x (or that P) $\Rightarrow$ S is conscious (a conscious being).

Dretske then connects creature and state consciousness:

4. S is conscious of x (or that P) $\Rightarrow$ S is in a conscious state of some sort.

The conscious state represents x or that P; it does not entail that x exists (or P is true).

Now Dretske is ready to pose two questions: Can one have conscious experiences without being conscious that one is having them? Can there be conscious states in a person who is not aware of them?

An affirmative answer to the first question is directly opposed to higher-order views. His test case shows that there can be differences in a person's conscious experiences (experienced differences require differences in experiences), which the person is not conscious of. One can be in conscious states without being transitively conscious of them. This challenges the higher-order thought theories: one can be conscious of things (as his tests indicate), entailing being in conscious states, without being fact aware of them, i.e. without having concept-involving states directed on to them. This also challenges the higher-order monitoring view, which holds that a state being conscious is one that is monitored; for what is the use of monitoring a state if it does not make one conscious of it?

So what is Dretske's answer to the problem of consciousness? In brief his view is that creatures like us are intransitively conscious not because we are aware *of* our mental states but because we are aware *with* our conscious states. We have intransitively conscious mental states because that is the way to be transitively conscious of things. And we need to be transitively conscious of things to survive.

RG 18.11.03

SS4 Consciousness

Seminar 8: The Concept of Consciousness

One central issue in the philosophy of consciousness is how the qualitative character of experience figures in any account of consciousness. Higher-order theories claim one can separate the issues of consciousness and qualia. Representationalists typically argue that qualia can be fully explained in terms of functional and cognitive properties. Ned Block has steadfastly argued of the phenomenal character of experience, e.g. the blueness of my visual experience of something blue, that it can only be explained by the presence of intrinsic qualities of experience (properties of the brain), irreducible to functional, intentional or cognitive properties. If qualia cannot be so explained away they would remain an inextricable part of the problem of consciousness. In this paper the view informs his argument that the term 'consciousness' is used in different ways and these should not be conflated.

Concepts of Consciousness

Block claims that 'consciousness' means several different things, that is to say the concept of consciousness is a mongrel concept. In particular he distinguishes between phenomenal consciousness (P-consciousness) and access consciousness (A-consciousness).

P-consciousness is that kind of consciousness which is constituted by phenomenally conscious properties, i.e. qualia. A state is P-conscious if there is something it is like for one to be in that state. Block claims that properties of P-consciousness may always be representational, but they are nevertheless distinct from 'cognitive, representational or functional properties'. In other words, the phenomenal properties exist prior to their representational functions. P-consciousness is the subject of the explanatory gap.

A-consciousness is the kind of consciousness which involves the accessibility of information for the purposes of reasoning and action. A-consciousness is constituted by the accessibility of representational states. A-consciousness is, for Block, an information processing correlate of P-consciousness (cp Baar's 'broadcasting in a global workspace' and Dennett's 'cerebral celebrity').

Block claims, against computationalists, that A- and P- consciousness are distinct and that the biological details are important. However, A- and P- consciousness can interact. The basic differences are: the content of P-consciousness is phenomenal (intransitive), P-consciousness is not a functional notion (within a module), and there are P-conscious types (e.g. sensations); the content of A-consciousness is representational (always transitive), A-consciousness is a functional notion, and token A-conscious states may not be A-conscious at other times (e.g. propositional attitudes).

Double-Disassociation of P- and A-consciousness

According to Block, the two kinds of consciousness are different because one can have either without the other.

Philosophers have noted how the lack of phenomenal consciousness produces a lack of information, therefore they have argued that phenomenal consciousness is amenable to analysis in functional terms, e.g. *blindsight*. Block agrees that in such cases of blindsight, A-consciousness is also lacking (poor access). But there are possible cases of *superblindsight* in which P-consciousness is lacking but A-consciousness (medium access) is present. A superblindsight person may have access to information without being phenomenally conscious (just knows that something is the case rather than knowing via a visual experience).

Conversely, one can be P-conscious without being A-conscious: the brain might be damaged such that phenomenally conscious states might not be A-conscious, or we might perceive noises without them being accessible.

Problems arise because we think of 'consciousness' as a univocal concept.

RG 25.11.03

SS4 Consciousness

Seminar 9: Dispositional Higher-Order Thought Model

Carruthers proceeds towards his preferred model of consciousness - a dispositional higher-order thought model - via consideration of a number of alternatives.

Some initial distinctions are noted: between creature consciousness and state conscious (due to Rosenthal); and between intransitive and transitive consciousness. In the case of transitive creature consciousness, however, it is problematic whether we say transitive creature consciousness requires being in a conscious state; if we say 'yes' we require an understanding of state consciousness, and saying 'no' seems paradoxical. The central distinction is between phenomenal consciousness and a functional notion (e.g. access consciousness). Carruthers favours the view that phenomenal consciousness can be explained in functional/representational terms. (He argues elsewhere that cognitive accounts of phenomenal consciousness are more likely to lead to explanatory success than neurological ones.)

The next issue is the type of representational account to give. Carruthers notes some of the advantages of first-order representational (FOR) theories, e.g. the transparency of experience. This is plausible over a range of perceptual experiences, including pain. Carruthers, however, objects that such externalist accounts of experiential content cannot account for 'Swampman' type objections. So he favours an internalist account of content.

The main objection to FOR theories is that, whilst they may explain the phenomenal character of the phenomenal world (worldly experience), they do not explain the phenomenal character of our experience (mental state subjectivity). To understand the latter kind of phenomenal properties, we need to understand how creatures can discriminate between their own experiences. HOTs are required for this. To meet the transparency objection HOT theory claims that all that gets added is an aspect of *seeming* to first-order content. This is the difference of content between the content 'red' and the content 'seeming red'.

The second objection for FOR theories is that they cannot explain the difference between conscious and non-conscious experience. Two unacceptable accounts are that (1) an experience is only conscious if it is available to the right sorts of decision making processes (Kirk), or (2) that an experience may be conscious but unavailable to a subject unless phenomenal concepts are applied to it (Tye).

Carruthers claims that phenomenal consciousness can be explained as occurring where perceptual information is made available to HOTs (recognitional capacities). This will also explain why we think there are qualia, i.e. non-relational properties of experience and the conceptual (rather than metaphysical) possibility of inverted qualia.

HOT theorists seem to have to accept that animals and young children, who do not have the appropriate HOTs about experiences *qua* experiences, will not be conscious in the way that we are. They would either have to have concepts of seeming or non-conceptual discriminatory capacities (HOE model). The HOE model is ruled out for reasons of computational complexity; also it is not required for binding and does not serve the purpose of a higher-order level of functioning, namely to distinguish seeming states.

HOTs can be given independent explanation in evolutionary terms: a theory of mind module might have evolved to help predict other people; this might then have been applied to the subjects' own thoughts. Carruthers favours a dispositional HOT theory, which responds to the problem of cognitive overload. A state is conscious if it is contained in a short-term memory which is then available to HOTs.

RG 2.12.03