ProtoSociology
An International Journal of Interdisciplinary Research

Volume 22, 2006
Compositionality, Concepts and Representations II:
New Problems in Cognitive Science

CONTENTS

D. Dennett's brand of anti-representationalism:
a key to philosophical issues of cognitive science ........................................ 5
Sofia Miguens

Externalist Thoughts and the Scope of Linguistics.................................... 23
Kent Johnson

Presentations and Symbols: What Cognition Requires of
Representationalism .................................................................................. 40
Christopher Viger

Perceptual Content.................................................................................... 60
Elka Shortsleeve and Kelly Trogdon

The Nonconceptual Contents of our Minds .......................................... 78
Diana I. Pérez

From Representation and Identification to Misrepresentation and
Misidentification ....................................................................................... 99
Urszula Żegleń

Personal-Subpersonal: The Problems of the Inter-level Relations .......... 120
Liza Skidelsky

VP-Ellipsis and the Case for Representationalism in Semantics ............ 140
Anne Bezuidenhout
ON CONTEMPORARY PHILOSOPHY

Science and Reality ................................................................. 171
Nicolas Rescher

Mixed Methods and Ontological Commitments ..................... 186
Steven Miller, Marcel Fredericks

Contributors .............................................................................. 207
Impressum .............................................................................. 209
On ProtoSociology .................................................................. 210
Published Volumes ................................................................. 211
Digital Volumes available ....................................................... 214
Bookpublications of the Project ............................................... 215
Cooperations – Announcements .............................................. 220
D. Dennett’s brand of anti-representationalism: a key to philosophical issues of cognitive science

Sofia Miguens

Abstract:
Although D. Dennett is sometimes accused of insensitivity to ‘real’, first-person problems of the mind, his Intentional Systems Theory offers a comprehensive, cognitive science grounded, account of the nature of subjectivity. This account involves views on intentionality (concerning the nature of the representation relation, content, psychological explanation), consciousness (comprising a functionalist model, a second order, belief-like, theory of self-awareness, and a deflationary view of qualia), personhood and freedom of action (concerning what must be in place in terms of cognition for the mentalistic concepts of ‘person’ and ‘action’ to apply). Since Dennett defends that the principles for understanding intentionality and consciousness are the same, in order to understand his brand of anti-representationalism we must deal with both intentionality and consciousness. That is what I will do in this article. I will also discuss the metaphysical implications of anti-representationalism, and in general use Dennett’s work as a key to describe how a range of philosophical issues of cognitive science appear from an anti-representationalist point of view.

“There are two major traditions in modern theorizing about the mind, one that we will call representationalist, and one that we’ll call eliminativist. Representationalists hold that postulating representational states is essential to the theory of cognition (...) Eliminativists, by contrast, think that psychological theories can dispense with such semantic notions as representation. According to eliminativists the appropriate vocabulary for psychological theorizing is neurological or perhaps behavioral.”

1. Cognitive science, mind and method
Cognitive science research bears on our understanding of fundamental questions about the mind, such as the nature of representation, consciousness and
action. Drawing on such research D. Dennett's work of the last thirty five years offers proposals about how to think about each of these issues. His views are admittedly controversial – he tends to be accused of insensitivity to the 'real', first-person, problems of the mind, but in fact he provides a comprehensive, cognitive science grounded, account of the nature of subjectivity.

For Dennett intentionality, and not consciousness, is the most basic issue about the mind, and the notion of interpreter is essential to deal with it (Dennett 1987). His theory of consciousness involves a critique of *qualia* and the Cartesian Theater, but also positive views, namely a funcionalist model aimed at explaining the status of the centredness, seriality and continuity of mental life and a belief-like, second-order, account of self-awareness (Dennett 1991).

It is from views on intentionality and consciousness – which will be explored in this article – that a practical philosophy arises, concerning what must be in place in terms of cognition for the mentalistic concepts of 'person' and 'action' to apply to cognitive systems (Dennett 1983, Dennett 2003).

Anti-representationalism is the core of Dennett's view of the mind. Anti-representationalism is the idea according to which for there to be mind there need be no such thing as real inner representations in those physical systems we take to be rational agents. For Dennett this means that there's no such thing as representations short of a stance, a strategy towards certain physical systems. Many people consider these formulation already unintelligible, or question begging. Critics in general took it as interpretivism (i.e. mind would depend on a merely instrumental attribution of mentality by an interpreter). In fact, in order to fully understand Dennett's brand of anti-representationalism we must see it as a more sophisticated view, including (i) the idea that mentalistic notions apply to agents globally considered and not to inner states, (ii) the fact that they apply depends on recognition of patterns of behaviour and also on agents inner design for recognition (which means questions of representation

---


2. Unlike many philosophers of mind (such as D. Chalmers, T. Nagel, or J. Searle) Dennett rejects the idea that consciousness is the fundamental problem about the mind: consciousness is a kind of representation, there's no such thing as 'phenomenal consciousness' as an extra property of the world. Dualism aside, this can be seen as a leibnizian conviction about the continuity of awareness and self-awareness, as opposed to the Cartesian idea of an absolute character of consciousness.

3. Whether it is dependent on an evolutionary or a transcendental argument (Fodor & Lepore 1992, 154).
and design should come together in the theory of mind\(^4\), (iii) a position about the metaphysics of cognition, which involves taking seriously the difference between vehicles of representation (e.g. neural events) and what is represented. The whole view may arguably push Dennett into denying metaphysical realism (Haugeland 1997: 267-304) but it is certainly not instrumentalist.

Before addressing Dennett's general positions I want to point out something that may help a reader navigate through his work. Given his dislike of tempting subjects such as qualia and zombies, and also his dislike of overlinguified and overargued approaches to questions of mind, Dennett sometimes seems to be very far from protocols of argument in philosophy, not so much a philosophers' philosopher as a philosopher for those outside of the discipline who are interested in philosophical issues. He casts a suspicious look on some ongoing debates in the philosophy of mind and often goes straight to cognitive science\(^5\). This is on the one hand a symptom of what Dennett calls his technophile inclinations - in his own words, if not for being born into a arts and humanities academic family he would probably have become an engineer (Self-Portrait\(^6\))- but we may also see it as a quinean, naturalized epistemology, imperative put to practice in terms of method, disrupting supposed boundaries between what should be scientific and what should be philosophic in the theory of mind. In consequence, Dennett's work often takes the form of an effort to relate cognitive science with phenomenological descriptions of mind (taking phenomenology to mean a neutral description of mental goings-on, at the personal level, the level we access our own minds\(^7\)). Bringing together this phenomenological take a cognitive system may have of itself with a third-person approach involves specific problems: cognitive science is involved in what Dennett calls (Dennett 1991) the dismantling of the witness protection program or a close-up of the observer. There's no witness, no observer inside a cognitive system which we take to be mental (ourselves included, of course) and still there's mind.

Yet, although taking cognitive science seriously involves trying to find out whether self-knowledge can be disrupted from a third-person point of view, this may in fact go together with the impossibility of being an eliminativist.

\(^4\) Naturally, this ended up focusing Dennett's interests on Darwinism (Dennett 1995).

\(^5\) Artificial Intelligence namely (e.g. the work of people such as A. Newell, M. Minsky, H. Simon), is abundantly referred and explicitly considered as a source of inspiration.

\(^6\) Dennett 1998, 356.

\(^7\) When we describe ourselves from within what we are doing is not forging hypothesis about the brain but rather phenomenology, in this sense.
And that is exactly the case with Dennett. Anyway if we consider that self-knowledge stops – it goes only as far as the mind goes, or rather, not even that far, and definitely does not reach cognitive architecture or physical workings – Dennett’s decision not to separate philosophical theory of mind from third-person approaches to cognition doesn’t seem such a bad one.

2. A few remarks about the origins of Intentional Systems Theory

Since Dennett’s anti-representationalism results from what was going on in philosophy in the 50s and 60s, I will take a step back to reformulate the question about Dennett’s anti-representationalism as a question about what makes for an antirepresentationalist such as Dennett. In epistemological and ontological questions, Dennett’s work was inspired by ideas of Ryle, Wittgenstein, Quine and Putnam. Ryle showed him that the alternative between materialist and idealist monism depended on an incorrect approach to the problem of mind, and Wittgenstein the groundlessness of any presumption of epistemic authority or direct access to the structure of the world in mental descriptions.

Wittgenstein and Ryle together provided an example of how description of mind at the personal level should go, and of how talk of mind should be taken, if not as authoritative introspective description of thinking substances. From Quine, Dennett took the naturalist outlook and the regard for natural science, and also (as Fodor put it) nihilism about meaning – the idea that ‘there are no semantic engines’. Putnam’s functionalism, a sophisticated kind of materialism, allowed him to articulate his opposition to psychophysical identity theory and gave him the idea of the mental as incorrigibility of personal level access. His

---

8 An approach which assumed the disjunction ‘either there are minds, or there are bodies’, before asking which is reducible to which.

9 Also the problem of the limit of mentalistic descriptions: for instance I say ‘2+2 =4’ / ‘How do you do that?’ / ‘I don’t know, I just know it’; still, the fact that the subject cannot answer the question doesn’t imply it can’t be answered. The story of Dennett and Wittgenstein is an interesting one: Dennett ended up in Oxford, where Ryle ‘had been king’, for his PhD, but he had actually set out to England in search of his philosophical hero, Wittgenstein. Yet, the way philosophy was practiced by self-professed wittgensteinians displeased him, mainly because it meant taking the legitimate difference between personal and subpersonal levels approaches to mind as an excuse not to care about subpersonal level questions. So Dennett ‘quit trying to be a wittgenteinian and simply took what he thought he learned from the Philosophical Investigations, trying to put it to work’ (Dennett 1998, 365).
thinking about action was influenced by Ch. Taylor's philosophical critique of psychological behaviourism, as an utopian search of a pure data language to explain behaviour, as well as by E. Anscombe's analysis of intentional action in *Intention*.

The starting point for the theory of mind is thus 'phenomenology', and that's what Wittgenstein (in the *Philosophical Investigations*) and Ryle (in *The Concept of Mind*) do: a report of mental happenings, whether these be feelings of pain or mathematical thoughts, without any presumption of explanation, or inquiry about physical basis of mind. Strangely enough, these philosophers, on whose work Dennett molds his own, are often taken as behaviorists. Yet, behaviourism, Skinner type of behaviourism, was exactly what stood opposed to the methodological revolution in psychology which prompted Dennett's philosophy, and which we can loosely call cognitivism, in the sense of informational-computational approach to the mind. Dennett has always opposed what he called 'peripherist behaviorism' and defended the indispensable role of intentional and teleological descriptions of behaviour – initially he called this position centralism (Dennett 1969), later Intentional Stance (Dennett 1987). Psychological and philosophical behaviorism have always been rather different anyway, and it is philosophical behaviorists Dennett takes as models. But there's one point where Dennett profoundly disagrees with a rylean-wittgensteinian approach to the mind: the idea that a self-sufficient conceptual analysis could be a proper philosophical method, which easily pairs up with contempt for natural science. What's important about Wittgenstein and Ryle is the fact that they try to conceive the status of the personal level – the level at which we describe our minds from within, and where pulleys and levers questions, i.e. mechanical considerations, are not legitimate. Such personal/subpersonal distinction is vital, when dealing with mind and cognition, to set apart what is descriptive and what is explanatory. Still, even if personal level approach should consist of phenomenological data-gathering for the theory of mind, this should be brought together with scientific research about cognition. Luckily, to help him step past the lack of interest in the philosophical bearing of science, he had brought from America the influence of Quine, and

10 In fact, this hides a deeper behaviourist commitment: Dennett thinks behaviourist principles should be applied to the inside of the brain and not the whole organism ('the skin is not such an important frontier') and translate into evolution by natural selection (Dennett 1978, Dennett 1995).

11 And, in the cases of Ryle and Wittgenstein, also for questions which excite philosophers of mind nowadays, such as physicalism.
the imperative, for a proper quinean, not to overlook science. This led Dennett to a hybrid philosophy / cognitive science style of approach, which sets him equally apart both from Ryle and Wittgenstein conceptual analysis as from much contemporary philosophy of mind.

Putnam's functionalism came as a final touch. In his manifests from the 60s, Putnam declared the mind-body problem to be not an empirical problem but a logical, linguistic one, with nothing to do with a supposedly unique character of human subjective experience, since it would arise for any system characterized by an asymmetry of access to its own physical and functional (logical) states. So from Putnam, Dennett took not only functionalism as a critique of identity theory but also a characterization of mind as incorrigible self-access to functional states.

The result is Intentional Systems Theory (IST henceforth): a moderate realism about the nature of representation (Dennett 1998, 95), going together with a teleofunctionalist view of content and the refusal to consider psychological explanation as explanation properly speaking. As for consciousness, the theory includes a deflationary view of (so called) qualia, and a second-order, belief-like, account of awareness, based on a functionalist model – the Multiple Drafts Model (Dennett 1991). Dennett's gradualist position about 'persons' and 'actions' (Dennett 1983, Dennett 2003) comes of these.

3. Representation: sententialism, eliminativism or interpretation

But what is anti-representationalism, if not straight Churchland-style eliminativism? In IST's case it results from (i) a quinean option for interpretation in the theory of mind (ultimately explored through the idea of real patterns

12 Cf. *Self-Portrait* «What do you get when you cross a Ryle with a Quine? A Dennett, apparently». (Dennett 1998, 365). From Quine Dennett also took the idea that a science of intention is not possible: for him – and he's not alone here – history of contemporary philosophy of mind starts with Quine's famous paragraph (§45) of *Word and Object* (1960) about the impossibility of a science of intention.

13 Such as *Minds and Machines* from 1960 and *The Nature of Mental States* from 1967.

14 Later, he also admitted to the influence of R. Rorty's papers from the 70s, where Rorty defended that what makes something mental is not the fact that it explains behavior or the fact that it is, or is not, a property of a physical entity, but the fact that certain reports about it have a certain status of incorrigibility. Rorty stressed then that incorrigibility is not infallibility: reports by a cognitive system about what goes on in it may well be wrong, still knowledge-claims cannot be overridden.
recognized from the Intentional Stance), (ii) darwinism (it is design that accounts for both patterns and recognition, and Darwinian principles of evolution by natural selection should account for the origin and nature of design). This eventually leads to a denial of metaphysical realism.

While a self-professed physicalism, IST's first concern is in fact the status of mentalistic language in the theory of cognition. It was this option that eventually led to the defense of a real patterns (as opposed to Language of Thought intentional realism) view of the nature of representation. Whenever the behaviour of some physical system is being understood or predicted mentalistically, the intentional stance is being adopted (according to Dennett, there are other stances we may take towards a physical system, namely the physical stance — the physicist's approach, and the design stance — the biologist, or computer programmer, focus on function)7. By taking the intentional stance a cognitive system makes sense of the behaviour of another without any idea of inner workings. How is it possible that minds read minds this way? It is possible because patterns of behaviour, relations between the system and those aspects of environment which matter to it, are being recognized by another cognitive system. Before seeing this interpretivism as untenably behavioristic and anti-naturalist we should consider that we needn't even be talking about humans here — what Dennett is saying is that we cannot avail ourselves with the notion of representation without considering relations between agents and their environment. What we, in the theory, call representations depends on interpretation in that it depends on recognizing, from the intentional stance, the embeddeness of behaviour of agents in the environment. It is important here that there are basic representations, which are action-oriented and about what is, for specific agents, relevant in the environment8. Contrary to what

15 There have been several versions of Dennett's Intentional Stance, from the rough instrumentality of Intentional Systems (Dennett 1978, 3) Dennett himself dismissed, to Real Patterns (Dennett 1998, 95). I will consider only this last formulation.

16 Dennett's physicalism is more like a general commitment than a development such as that carried on by authors such F. Jackson, D. Chalmers ou J. Kim. It boils down to being a quinean and giving priority (as he formulates it in Content and Consciousness) to the language of natural science, avoiding a kind of 'double speak'.

17 The physical stance is the most general of the three — it applies to everything there is. Still, according to IST intentional level entities cannot be reduced to this lower level (that's the reality of patterns). There's no clear frontier between those systems predictable from the Intentional Stance, and those that are not, and the Intentional Stance does not allow us to distinguish true believers from systems merely predictable as if they were believers, but for Dennett this is no insufficiency — it simply reflects the way the world is.

18 All this is close to authors such as A. Clark and R. Millikan. To grasp what's characteristic
instrumentalist readings assume, Dennett thinks that an interpreter (until now, a mind-recognizer) does not arbitrarily create intentional interpretations of other cognitive systems, rather representation talk is made legitimate by recognition of environment-embedded patterns of behaviour. This should be accounted for as a matter of design-for-representation-using and and representation-recognizing and inevitably it leads from the question of representation to that of the nature of design. Thus, Dennett has to come clear with his view of design: is it real or not? To sum up, he thinks design is real in that it results from an algorithmic, mindless process of evolution by natural selection (Dennett 1995), but depends on interpretation in that functions – and functions of inner design-for-representation included – depend on recognition to be taken as specific functions\(^{19}\).

So although representation relations depend on recognition, they do subsist and more must be said about their content. Dennett does this in terms of a tele-functionalist account: content is not within, where vehicles of representation occur, it is a function of function, externally determined. But what is external? If action-oriented representations are the basic type of representations, the world is represented differently by different cognitive systems. For instance, different animal perceptive systems evolved to respond to the environment’s affordances, so different cognitive systems will be different world-makers (taking well known examples from the literature, there will be worlds in which location-of-nectar is especially salient, worlds where that’s the case with possible-nutrition-which-flies-quickly-through-visual-field, or toxic-surface-water-to-be-avoided, etc). This Dennett calls notional worlds, a concept that does not apply to vehicles of representation within a cognitive system, nor to the world directly, but rather to what is represented\(^ {20}\).

---

19 This is why his teleofunctionalist view of content differs from the realism of people such as Millikan or F. Dretske. Dennett’s way of posing the problem of nature and content of representation is akin to the kantian problem of status of perception of teleology in nature once natural science banishes it. Once this problem is made more specific through cognitive science, what we have are cognitive systems which somehow recognize mentality in nature and are which themselves (at least the biological ones) the result of evolution in nature. This is one way to formulate the problem cognition poses for metaphysics.

20 Cf. Evolution Error and Intentionality (Dennett 1987, 87). There are several problems with this externalism-plus-notional-worlds view. Still, two important traits of the view are: opposition (i) to full determination of content – eventually of semantic intentions in humans – and (ii) to intrinsic intentionality.
What do these views about nature and content of representations amount to when it comes to giving an account of psychological explanation? Basically, they keep Dennett from subscribing to anything Fodor's hard-core representationalist realism buys him: causal thinking about the intentional level of reality, psychological explanation as nomological, and so proper explanation – the whole constellation of special sciences, \textit{ceteris paribus} laws, independence of psychological generalizations from physical implementation. But for Dennett, Fodor's sentences-in-the-head view conflates intentionality with design: being an engineering hypothesis Language of Thought commits the rylean sin of taking a conceptual answer to be a causal answer. The representational-computational theory of mind takes notions (language, representation) which apply to the whole cognitive system as if they could do explanatory work at the subpersonal level. For Dennett this amounts to letting a ghostly central observer sneak in in the theory. Plus, Dennett sees no reason to accept the common design (Dennett 1978: 90) of cognitive systems Language of Thought implies – for him subpersonal cognitive psychology is just a theory of implementation of that which, intentionally interpreted, are representations and computations.

The next move is to clarify the relations between folk psychology, (scientific) subpersonal cognitive psychology and IST as a philosophical view of cognition (Dennett 1987, 43). And the way he sees it, there is no reduction of intentional characterizations to physical descriptions – but also no explanatory power of mentalistic notions.

Not only did the relation between representation and design lead Dennett to Darwinism but also to what he calls, in \textit{Kinds of Minds} (Dennett 1996) an evolutionary point of view on ontology and metaphysics. From an evolutionary perspective there's a plurality of minds. But can pluralism be defended together with physicalism, understood as priority given to the physical stance, and a constitutive reference to physics, as basic science, when the question is what fundamentally exists? Again, the question of reduction comes up. There are several ways to be an anti-reductionist about the mind (i.e. to deny that the intentional may be reduced to entities and laws situated below within the framework of a hierarchical conception of nature), the main ones represented in the philosophy of mind by Fodor and Davidson. To understand Dennett's anti-reductionism we must focus not on the situation where a human instrumentally interprets another as mental, but

\footnote{I'm assuming Fodor's views about an intrinsically representational level of reality define the field and the problem.}
rather on the status of *what is represented*, and on the concept of an interpreter.

What is the ‘interpreter’? Answering this question implies clearly stating the metaphysical implications of IST. J. Haugeland does that in an illuminating way, bringing IST together with an idea which is actually one of its main targets: J. Searle’s concept of intrinsic intentionality – the idea that one can legitimately distinguish real believers from which are merely predictable from the intentional stance. Haugeland suggests that we use searlean intrinsic intentionality and Dennett’s interpreter to clarify one another. Both Dennett and Searle claim that intentionality is *normative*, and Searle proposes the notion of aspectual shape (of anything mental) to say what is meant by normativity. Now, whereas Searle simply characterizes aspectual shape, and cannot account for it, Dennett, with his idea of interpreter, can. Let’s formulate the problem as ‘how can physical configurations of brains resulting from evolution by natural selection be intrinsically normative’, and take a game of chess as an example of taking something as something. How does a player see that a horse threatens a pawn? No causal history accounts for the perception of horse-as-horse or pawn-as-pawn: only a stance, a constitutive commitment, from which depends the very reality of chess phenomena, does. These are not identifiable with physical pieces of ivory (or patterns on a screen, or marks on a paper); rather they are constituted by standards. Whose standards? The subject’s – and what Dennett is saying about representation is that the constitutive role of interpretation stands to mentality as the subject stands to the reality of chess phenomena.

Haugeland sees the intentional stance view of representation as kantian constitution: objects of perception, action or thought, are understandable as what they are only in terms of a commitment of a subjectivity-qua-unity (i.e. the interpreter). It is this interpreter we need to understand in order to have a metaphysics of representation. The interpreter is neither a place in the brain, nor a self, it’s more like a *taking-as* function. This normativity is a feature of subjectivity which for instance a neurobiological theory of consciousness can-

---

22 Haugeland 1997, *Understanding Dennett and Searle and Pattern and Being*.
23 J. Searle’s Chinese-Room mental experiment (cf. his 1980 * Minds Brains and Programs*) is meant as a critique of attributed, as-if, intentionality, and a defense of genuine intentionality.
24 *What is represented* is represented as such and such, as an aspect – that’s what distinguishes a neuron from something mental like a memory, both being, at an instant t, unconscious (Searle 1992).
25 This is what Dennett thinks of as the reality of patterns. Another example Dennett always uses is the Game of Life (Dennett 1998, 105-110)
not account for, nor should, because it is not by itself a theory of representation, but of the vehicles of representation. The chess analogy is useful to dispel any suspicion about the mysterious character of normativity – what is there is a stance, without which there’s neither subjectivity nor objecthood: that’s the two-sided nature of mind as representation, necessary for searlean aspectual shape which distinguishes anything mental as mental. Through this we can understand searlean intrinsic intentionality as subjectivity to itself, or self-mental-interpretation, and also see the other side of the problem of representation as what B. Cantwell Smith (Cantwell Smith 1996) calls the ‘origin of objects’. Cantwell Smith points out that physics does not buy us objects – individuated, continuous through time, discrete and identifiable represented objects – but only particularity and locality. Physicalist theories of mind should take into consideration, instead of assuming existence identity or persistence across time of intentional objects. Being a represented object is not a local property of a space-time region: objects (individuation in what is represented) are there only for certain entities. Object is not a causal notion, but an historical concept, and that’s the main reason why representational notions should not be taken as physically effective.

Haugeland and Cantwell Smith help shaping Dennett’s anti-representationalism and also his view of psychological explanation: anti-representationalism is not a view about human agents’ capacity for interpreting one another by means of strategically attributing mental states, but a view about the nature and status of representation in the world according to which it is only at the level of a subject / object interface that ontologies (Dennett’s notional worlds) are defined (even if it may still be claimed that the world is metaphysically one). So, in spite of accusations of behaviourism and instrumentalism, IST in fact involves the irreducibility of subjectivity – not of qualia but of this essential role of an interpreter for representation. This also means there are at least two different unities a theory of mind must account for: the interpreter, the unity involved in intentionality, and another unity, a represented unity in a mind, for itself – the self, which, from a cognitive point of view, Dennett sees as virtual unification in a multiple agents cognitive system. The first is a metaphysical question about cognition; the latter is a question about the individual subject’s subpersonal mechanisms of cognition, to be answered in terms of brain\(^2\). This

---

\(^2\) Dennett himself sometimes speaks of notional worlds as sets of subjects’ reports, and even hints at comparing subjects’ reports of their beliefs with events in their brains (this would be true reality). I think that amounts to denying his own best intuitions that the brain is not thought, representation or understanding – in order to understand thought what we must
is crucial for the next step in the theory of mind: a theory of consciousness, partly built on this view of representation, but also involving the proposal of a cognitive model, and a set of epistemological considerations about qualia. The fact that Dennett defends a representationalist theory of consciousness (i.e. consciousness is a form of representation, principles for understanding intentionality and consciousness are the same) makes it necessary, in order to understand his anti-representationalism, to deal with intentionality and consciousness together.

4. Consciousness as an imprecise Cartesian Theater

Dennett’s theory of consciousness is bound to frustrate fans of phenomenal consciousness, *qualia* and *zombies*: it is a deflationary theory, in which consciousness is explained in terms of representation, self-access, and incorrigibility of self-reports (in fact, *qualia* and *zombies* don’t even make sense within a representational theory of consciousness). Dennett’s theory of consciousness deals not with the so-called hard problem of consciousness, but with a set of problems (kinds of access, unification of a stream of consciousness from multiple drafts at the subpersonal level, self-representation and virtual center, self-reports, the difference language makes in a mind, etc). The type of solutions Dennett comes up with make it obvious that what he is offering is at least partly a cognitive model. This earned him the accusation of offering no arguments and no answer to the metaphysical problem of phenomenal consciousness. For Dennett, though, this is how a theory of consciousness should be, if it does not go after bogey-problems.

Dennett’s theory of consciousness is developed mainly through the Multiples Drafts Model (put forward in the 90s, in *Consciousness Explained*, as a successor of the *Brainstorms* model). Funcionalist models (which are, in spite of the proclaimed opposition to the Cartesian Theater27, the starting point for a belief-like theory of inner awareness) are the constructive part of Dennett’s theory of consciousness — through them Dennett aims at filling the (supposed)

understand first is representation and not neurophysiology or cognitive architecture, and that in order to understand representation we must understand the interpreter and the origin of objects.

27 The Cartesian Theater — involving a center where everything comes together to be observed by the mind — is a widely endorsed (and not necessarily dualist) metaphorical view of consciousness.
abyss between physiology and phenomenology. Plus, they allow him to deal with empirical issues of cognitive science, such as the status of mental images and dreams, production of language, blindsight, temporal anomalies (such as color phi-phenomenon and Libet’s problems of backward referral in time and delay of consciousness of intention) (Dennett 1992). The destructive, and better known, part of his theory of consciousness is an argument for quining qualia—decisive for rejecting the usual ways of vindicating phenomenal consciousness: zombies, knowledge argument, what-it’s-like.

How does one go from a cognitive model to the elimination of *qualia*? Unlike what happens when dealing with intentionality (the agent’s unity is simply assumed to consider behaviour and adaption to the environment), dealing with consciousness involves *zooming into* cognitive systems, characterized at subpersonal level(s) as (i) hardware-level parallel distributed processing, (ii) functional level competing agents. The assumption of a unity or center as given goes, and the unified representation of the system by the system becomes part of what is to be explained. Dennett’s critique of *qualia* is simply a matter of exploring the epistemological consequences of this absence of a real center where everything comes together for an observer—the Cartesian Theater metaphor for the mind. Still, even if the close-up of the observer reveals the absence of a real center, our own mind appears to each of us as an (imprecise) Cartesian Theater, and that is what the Multiple Drafts Model (MDM) aims at explaining. This means formulating subpersonal level hypothesis about what makes conscious experience unified, serial and continuous and also about what makes for our sense of control of intention. According to the MDM, unification depends on a user’s illusion of a virtual serial machine implemented on parallel distributed processing hardware (the brain of the body the self will call its own). Virtual is obviously not unreal—the mental life of the system is very much real for the system itself—and Dennett is not saying consciousness is an illusion (the illusion would be to think of the unified self as a separate substance, an out of the physical world pure experiencing) but trying to account in terms of cognitive architecture for the way we appear to ourselves. Self-representation is another necessary condition for the sense of being owner and author of phenomenology, and for the unification and centralization of control and intention. The

28 Of course the big problem is how to make subpersonal issues dealt with here fit with theories of interpretation, which assume the unity of the interpreter, and the interpretive status of representations, a notion freely used in the theory of consciousness.
29 Dennett borrows from M. Minsky these concept of illusion and also of virtuality.
30 Or rather multileveled self-representations, only some of them mental.
need for inner communication pushes for a represented unity (in language using humans, eventually an 'I'). This centralized virtual arena for unified control of behavior is necessary for what phenomenologically is feeling a self from within. Still, unified intention and initiation are also user's illusions in the sense that the 'control-component' (as the Brainstorms model has it) is replaceable with autonomous agents which can be 'dismissed' (i.e. they are bottom-up explainable). In such conditions, disunity and break-down lurk – the various pathologies of identity which have always interested those interested in mind (from schizophrenia and paranoia to multiple personality disorder and split brain) reflect this absence of a real center.

Not only centredness andseriality but also continuity of phenomenology has to be explained: part of Dennett's exploration of consciousness concerns the role of neglect and filling in for what may not be there in experience (from memory, to dreams, to perception) and his basic thesis is that if there's no observer there's no need for filling in. More specifically, awareness of x cannot exist without a belief and it is by principle impossible for the subject to decide, in what concerns his awareness of x, between appearance and reality (this may be called first-person verificationism). The blind spot, in vision, is emblematic of the kind of non-problematic neglect (we might call it normal anosognosia) Dennett is after and thinks is significant. In where there are no agents anxious for information, there's no problem if no information arrives. Dennett generalizes this to our feeling of our own phenomenology.

Having replaced the observer with specialized agents and identified awareness with availability, in memory, of knowledge—that is essential to understand what is probably the main test-case for MDM and first-person verificationism: the supposed stalinesque and orwellian revisions Dennett tests on colour phi-phenomenon and several other examples (Dennett 1991). It was already apparent in the Brainstorms model that memory was vital for Dennett's conception of consciousness. Personal level access of the system to itself, which is ultimately what Dennett calls consciousness, depends on different types of

31 The 'I' comes after a bodily based sense of self and it will be the reference for a very useful fiction for human agents, the narrative of personal identity (Dennett's metaphor for the I is that of a 'center of narrative gravity' (Dennett 1991).

32 Verificationist principles were first stated in the context of the analysis of dreams (Are Dreams Experiences?) in Brainstorms.

33 Another interesting idea about the epistemic prospects of introspection Dennett explores is the following: we, the users of our own brain, cannot distinguish what has always been there and what needs to be produced when a question is posed to agents. M. Minsky calls the conviction that we are able to make this distinction from within' the 'Imanence illusion'.
D. Dennett’s brand of anti-representationalism

subpersonal accesses. Namely expression, i.e. publication of any mental content, even to oneself (e.g. thinking linguistically), involves subpersonal transit between memory and mechanisms for reportability. If content of awareness is everything registered in memory, it may decay or be interfered with before it is recovered for public access. If ‘public’ should be identified with personal and includes the linguistic access the system has to itself, this determines the epistemic prospects of introspective authority, since what we call semantic intentions are based on the connection between control and reportability. Given the absence of difference between processing and status of contents of which we’re aware, consciousness depends on probes – it is self-monitorization which make consciousness as inner-awareness possible. This is developed through an higher-order view of consciousness (Dennett 1991:303). A system is conscious if and only if higher-order mental states, mental states which are about other mental states of the system, occur. And these mental states are, according to Dennett, beliefs, not perceptions34.

It is based on these ideas about how mental life appears to conscious beings such as ourselves that Dennet discusses and quines (i.e. eliminates) qualia. For him, the term names intrinsic, ineffable, private, uncorrigibly known characteristics of our mental lives. With qualia so defined, the claim that there are no qualia, amounts to saying that there could not be, in a subject’s mental experience, something that is immediately known and ineffable. There’s no real seem, no pure non-conceptual appearing (Dennett 1991, 369-411) – there simply cannot be any differences in conscious experience that subjects are not explicitly aware of35.

What follows from this? Dennett never says cognitive systems such as ourselves (in which there are no qualia) are not conscious – in fact, the very object of a theory of consciousness are those systems, among intentional systems, 34 A further qualification makes this a dispositionalist theory: first-order cognitive states (for instance, an instant of visual perception) could bring about many beliefs; these beliefs do not have to be explicit – in order to assure the possibility of conscious cognitive states it is sufficient that first-order cognitive states are kept in the system’s memory so that they are available to be the object of belief.

35 Three possible lines of criticism to this theory of consciousness are: (i) If we see it as the re-enactment of a classical wittgensteinian position Denett is not considering a problem Wittgenstein saw: where there’s no (inner space of) observation, there isn’t, strictly speaking, place for verificationism. Phenomenology is all there is, there’s no knowledge or observation of it, and only if there were, could qualia be considered a matter of incorrigibility; (ii) the whole thing is built on an equivocal reading of ‘seem’ (as appearing and judging) (D. Chalmers) and begs the question of phenomenal consciousness, (iii) what rules out non-epistemic (in F. Dretske’s sense) inner perception?.
which are conscious. Dennett theory of consciousness applies exclusively to humans – other animals, or even babies, simply are not subject to the Cartesian Theater illusion: those are minds less centralized that adult, linguistic, human minds, and not self-aware36. So even if Dennett does away with any aprioristic distinction between natural and artificial and intends to have us look at cognitive performances of actual, biologically based, minds as situated, with other (biological or not) possible kinds of minds, in the same design space (Dennett 1995), he also proposes criteria for human minds as a specific kind of minds. Cognitive architectures for communication and language make for thresholds: there are characteristics the minds of linguistic creatures possess because they are linguistic. Allowing namely for mental acts such as endorsing and affirming one’s own beliefs and making voluntary decisions, language simply makes human minds much more powerful than those of any other animals – for Dennett although this is not a difference in kind, only a difference in degree, it is still big enough to make all the difference in moral terms37.

36 In fact, Dennett’s approach of animal minds (Intentional Systems in Cognitive Ethology, Dennett 1987: 237) reinforces his anti-sentencialism: a philosopher’s task in cognitive ethology consists in contesting overattributions of consciousness and communication to other animals.

37 It is in his approach to personhood and freedom that Dennett deals with the specificity of human minds more directly. I will briefly sketch it. In the theory of action and moral philosophy the unity and continuity of agents and the possibility of deliberation and voluntary action are frequently taken for granted. But in fact humans are cognitive agents, not semantic engines with selves as real centers, so Dennett’s practical philosophy focuses on what must be in place in terms of cognition for the mentalistic concepts of person and freedom of action to apply. What comes of a route which takes him from an engineering approach to agents to an analysis of the fundamental concepts of morality is a gradualist, anti-essentialist view of personhood and a compatibilist account of freedom. According to the idea of conditions of personhood, ‘person’ is not an absolute concept, i.e conditions concerning intentionality, consciousness, language-dependent recognition of other minds through communication have to be fullfilled by a cognitive system for it to apply (Dennett 1978, Conditions of Personhood). Also, the concept of a person as a conscious, rational agent is not separable from the idea of an agent responsible for her own actions. This means that for Dennett, personhood is essential for full-blown consciousness in humans, and so – rather than being identifiable as sentience or qualia – consciousness depends on rational agency. Understanding freedom of action is thus essential for understanding full-blown consciousness in human minds, and understanding freedom involves understanding how the design for meta-reflexive self-control in agents came to be possible through evolution (Dennett 1983, Dennett 2003). Since free will worth wanting is neither noumenic freedom nor physical indeterminism – questions concerning freedom of action (responsability, guilt, praiseworthiness) are not micro-level questions, but macro-level questions involving mentality – discussions about physical determinism are off the point. What we call ‘freedom’ evolved in finite, designed, physical systems, dealing with incomplete information in a
7. Conclusion

The challenge in the study of cognition is whether it allows us to think of the world as ultimately mindless and Dennett's philosophy is an answer to that challenge. It is, thus, especially significant that his anti-representationalism does not in any way amount to an elimination of subjectivity from our conception of the world. In fact, it is remarkable that he has been accused of being something like a third-person absolutist, insensitive to 'real', first-person problems of mind, when in fact the central thesis of IST concerns the indispensability of mentalism. It is one thing to keep us from an appeal to intuition and to something unique about our kind of minds, and to argue that frontiers between mental and non-mental may be fuzzier, and the transition more gradual, than what we would intuitively believe, another thing to think we could ever do away with semantic notions when dealing with cognition.

Bibliography


deterministic world, as meta-reflective self-control, and to understand it we need (Dennett 2003): (i) a clarification of misconceptions of determinism as involving inevitability, (ii) the possibility, given the elbow room opened by 'evitability', of control in a deterministic world, through prevision of patterns (the capacity to smear over micro details is the key of 'evitability'), (iii) the distinction between control and avoidance (which are what we really want) from (supposed) noumenic freedom, (iv) the role of a linguistic self-representation (the 'I') in self-control, (v) an analysis of deliberation, from 'free floating rationales' in nature to full bloodied reasons used by human agents to control behavior. From the above Dennett gets the cognitive foundations for of an impersonal view of personhood. Yet that is not all that is needed for a person to be precisely him or herself – for that we have yet to consider self-creation as evaluation, through an analysis of the agent's commitment to processes of his/her own mental life. For this Dennett appeals to Ch. Taylor's and H. Frankfurt's conceptions of strong evaluation (evaluation of oneself and one's beliefs and desires, and not only of one's actions) and second-order volitions to capture this further dimension of self-control.

If we were absolute persons, all these questions would be practically redundant – but that is not the case. Thus practical questions make a powerful case for why we cannot do away with semantic notions: we cannot do away with them without doing away with ourselves.


