

## II

**Dretske, Fred** : *NATURALISING THE MIND*, (A Bradford Book - Jean Nicod Series), Cambridge, MA, 1995, The MIT Press, pp. 232.

This book\* is an extremely important contribution to the ongoing debate during the last two or three decades of our philosophically rich and pregnant but fast vanishing Century, whose story is a story of passage from non-naturalism to naturalism. While Wittgenstein offered therapeutic naturalised semantics to promote philosophical naturalism, Quine advocated naturalised epistemology to promote naturalised skepticism. In the philosophy of mind, however, the qualitative, the phenomenal, the what-it-is-like aspects of mind had defied all traditional and modern forms of naturalism. It was therefore an important question : How can the perplexing and baffling problems of phenomenal experience be accounted for? To answer this question Dretske offers in this book a provocative discourse to argue that in order to understand human mind, recourse to comprehend merely biological machinery that enables the mind to do its job is not enough. It's necessary to understand what the mind's job is and how it can be performed by a physical system, the nervous apparatus. This understanding is developed by Dretske within the framework of naturalism by claiming that the phenomenal aspects of perceptual experiences are one and the same as external, real world properties that experience represents world as having. He presents to us a completely naturalistic account of phenomenal consciousness, a theory which may be named as 'Representational Naturalism.'

The notion of representation is central to cognitive science. It is paradigmatic to current philosophy of science (R. Cummins, 1989). Within cognitive science, there are opposite camps, the representationists, who hold that internal representations exhibit a 'read-write-copy' linguistic structure, as well as the anti-representationists. According to critics like Andy Clark (1994),

---

\* The original inspiration to write this review comes from Professor S. V. Bokil's contribution to the Seminar on 'Critical Theory and Post-Modernism' (1997) held at the RIAS, Uni. of Madras. Bokil's thesis was : on matter what post modernism means, naturalism comes to stay. He encouraged me to write this modest summary.

anti-representationists are not as sceptical of the notion of representation, as they appear to be. They need a notion of deep representation, and they avow to identify mental representations with the 'cut-and-paste' neural architecture. They work with vector coding in a high dimensional space. Such representations are at the micro-level and may not after all require language. They are language-less. There arises, therefore, a dispute as to whether sentential or non-sentential paradigm is better suited for cognitive science. No decision to favour one or other is imminent because both use symbol manipulation and then confuse the implicit and explicit levels of representations, but in the reverse directions. Besides there are revisionary representationists, who favour a variation of radical, dynamic, and non-computational models. In between there are representationists who couple representations with segments of environment. All these notions probably lie in a continuum.

There is no one viable theory that holds sway. The whole story began with Jerry Fodor's 'strong' representational theory of mind that found representations of thought in mentalese (language of thought). Among the 'dizzying range of options' (Stephen Stich, 1992), one finds weak syntactic theory, causal co-variation theories (Dretske, 1988; Fodor 1987; and Fodor 1990), teleological (Millikan, 1984; Papineau, 1987), narrow theories (Devitt, 1990) and wide theories (Burge, 1979). The mainstay of the projects is to naturalise intentional categories, such as belief, desire etc. The crucial issue here is to know whether representation is a naturalistic phenomenon and it can be said to have a physicalistic or reductionistic character.

Stephen Stich cautions us by telling us that there is no one single project but a family of projects, much of which are located as interdisciplinary in its core. From this, he was led to theorise that there is a plurality of theories. From a common sense of folk psychological point of view, a distinction can be drawn between mental representations that are of the form I believe  $p$  (monadic structure) and I believe that  $p$  (dyadic structure). The monadic structure view entails a modularity assumption about the mind, according to which, it occurs within a module in an encapsulated form, thus warranting a hard-and-fast distinction between perception and cognition (inferential). This is shown to be a way of overcoming the substitutivity argument. The dyadic view, on the other hand, cannot see perception and cognition (conceptual) as integral forms of a

cognitive organism. An offshoot of the monadic view holds that I believe  $p$  is just analogous to I say/express  $p$  (where  $p$  is put within a scare quote) which questions the rationale of philosophical theories, which aim to give necessary and sufficient conditions for

### **Mental state $M$ has content $p$**

For them, mental state  $M$  has content  $p = p$ , and propose to offer a syntactic theory of mind (Stich, 1983). Fodor, on the other hand, deems it convenient to explore only sufficient conditions throughout his writings. The answer to the above question comes in the way by holding that mental representation is a naturalistic phenomenon. Robert Cummins (1989), for example, identifies such a project with a computational theory of cognition. Like Fodor, Cummins also seek sufficient conditions. The difference between Fodor and Cummins, in this regard, is significant in that while Fodor proves it in support of *ceteris paribus* laws, Cummins wants to pose the question whether the idea of mental representation plays the same role both in computational psychology. Fodor answers 'Yes', and Cummins answers 'No'. Cummins opens the way for computational states that are realised only indirectly rather than directly. Following the lead of Cummins, Terence Horgan was led in the direction of what I call an anti-Fodorian (he calls it as non-Fodorian, however) thesis : taking sufficient conditions as tractably sufficient conditions or surveyable, or as he prefers to put it, psycho- tectonically (a term used by Colin McGinn) realisable, he was led to formulate an argument against type-type identity of mental and computational states, finally arriving at a conclusion which states that it is not possible to specify tractably sufficient conditions. Horgan justifies the above by advancing a distinction between direct and indirect realisable conditions. The latter denies any strict identity between computational states and wetware realisation states. They are not computationally tractable because they are non-unique, and baroque (what functionalists call as multiply realisable, which provides an argument against type-type psycho-physical identification). The corresponding argument is called type-type identity of mental and computational states.

There are at least three crucial premises that led toward the above

standpoints. They are : there cannot be any one-to-one translation between different computational relations with two individuals or even within one and the same person. And hence there may be different computational relations between what is psycho-tectonically realised in human and what is psycho-tectonically realised in Martian Mentalese. It is therefore chauvinistic to hold that mental states can be individuated on account of their identity with syntactic or computational or realisation states. Hence, meaning can only be given in terms of broad content. I take this to be a significant improvement on Fodor, who overcomes the Twin-Earth problem by holding that translations are determinate. In contrast, Horgan uses a major premise of indeterminacy of translation, which looks like an apparent contradiction at the surface level. It appeals to me as a significant move for studying the very idea of representation (a naturalism without tractably specifiable sufficient conditions) which accommodates misrepresentation as well, which requires a similar fundamental premise saying that both experience and thought are representational in different ways (Dretske, 1985). This becomes the staple-cheese in Dretske's refreshingly information-theoretic approach, which bears a certain contrast with Fodor's own. If what I say is correct, then there is a broad agreement between Horgan's and Dretske's approaches.

So, the orthodox variety recommends a classical computational theory. The un-orthodox deal with cognitive or neural architecture. The former is sometimes identified with a linguistic paradigm, and the latter with a non-linguistic one, and it is supported by a connectionist paradigm. A connectionist paradigm uses parallel data processing for a design of the mental architecture. Stich expresses the scepticism as follows : there is no unique correct framework for theories in cognitive science (253).

Computational system is either individualistic or anti-individualistic (Putnam, Burge). Cummins has reason to take the latter side on account of the above distinction between direct and indirect realisation since beliefs and desires cannot be individuated independent of the environment, our common-sense psychology is said to be anti-individualistic, whereas scientific psychology cannot have that option: it should explain how beliefs are individuated. Externalism provides an option, but it acts more like a safety valve. It is so for the simple reason that naturalism demands coupling of cognitive organism with its

environment. Eliminativists who try to eliminate the folk elements have at least three options. First, they may argue that, following the above taxonomy of individualism, and anti-individualism, the ontologies of computational theories are different or incompatible with ontology of folk theory. However, the compatibilist position dominates the scene (Sterelny, 1990). Secondly, they may put up a case by holding that there are some features of scientific psychology which are not endorsed by folk psychology (Stich, 1983). Thirdly, we can argue that science approximates to the truth than folk: we have only the mad sort of thing called pan-eliminativism. Probably, Churchland endorses this: what is indeterminate in one theory is determinate in a future theory.

In all these, theorists use a modicum of philosophy of language for supporting philosophical conclusions. They use description theory, pure reference theory, and causal-historical theory. Descriptive theories argue that we can eliminate by description, pure referential theories seek a strong identity of the mental, making a case against belief psychology, causal-historicists advocate an externalist approach. The descriptivist theory is trivially true, whereas the causal-historical theorist is trivially false, and we are left with one option of pure reference. The trouble about these theories is that there is no one unique theory, and hence pluralism is the result. Consequently, even so, a weaker theory that identifies folk and science in a weaker sense (its strong sense in which eliminativism can never become true) suffers a casualty. Pluralism in the referential or descriptive theories is replaced by pluralism about externalism. There is a guarantee at least one or other versions may be true. The significance of the critique of the earlier-mentioned pluralism is that all such naturalistic projects suffer from a single defect that is germane to crossreferential inter-theoretic relationship (Stich). One may read this as providing scope for an argument against type-type identity of mental and computational states (Horgan). The same point can be said about the relation between linguistic version that translates into mentalese and the non-linguistic which fixes it in the cognitive architecture. Are they independent projects in the ultimate run? The answer is not so much clear.

Bald naturalism (R. Bernsterin's term) is completely out of question, and hence is chauvinistic. Non-reductive materialism has a promise, but no definite answer has emerged as yet. Versions of this are available from Davidson onwards

(Simon Evnine, 1991) up to Horgan who considers it as the third important paradigm, other than eliminativism and physicalism. In Horgan, it occurs in a very developed form which brings together tractably specifiable, and not tractably specifiable at the implementational level. From Horgan's point of view, if Cummin's account has anything to say on this matter, it is that representation theories fail because they can not take into consideration misrepresentation. Thus, misrepresentation becomes a fundamental issue. So it transpires that we need a more generic notion than representation. In fact, Cummins suggests such a more fundamental notion of *s*-representation, which is non-unique, but even then he could not succeed, by taking mental representation as a special case. Where actually this fails is that it takes seriously tractably specifiable condition for *s*-representation. Terence Horgan (1992) changes this into tractable non-specifiable conditions. Horgan argues that even if, they are tractable, they have no prospects of being physical and he was led towards a non-reductive physicalism. Spliced with all revisionary potency of supervenience (mental states are supervenient on physical states), the prospects look brighter. As if to rectify the very notion, Dretske calls attention to the generic status of misrepresentation. For Dretske, the very notion of representation must be rich enough to include misrepresentation as well. How to read the significance of this rich enough notion, which is claimed to be poised enough to meet even a sceptic about representation? Here lies the importance of Dretske's contribution towards the naturalising the mind, as I read it, which combines many of the positive features mentioned above.

The book under review is the latest in the series and it forms the 1994 Jean Nicod Lecture, in memory of the famous French logician Nicod, and Fred Dretske is the successor to this position. Dretske overcomes the unnecessary distinction between perception and cognition made by Fodor by replacing it with facts about representation and representational facts. The latter are mental facts or physically localisable facts in the brain. But facts about thoughts. Whereas Fodor holds that all mental representations are cognitively penetrable, dretske denies exactly this: some are, therefore, are not. In fact, they offer contrary solutions to the problem of misrepresentation (Kanthamani, 1998). He arrives at a hybrid notion of representation (26) that allows a generic status to misrepresentation. Granted that all mental facts are representational facts, all representational facts are informational states, one has to contend with facts about

representation which are external. So, we have to add the premise that the representational theory is an externalist theory of the mind. Coming as they are from external sources, it becomes incumbent to annex them to the representational facts. Dretske is obliged to tell us about introspection or self-knowledge, qualia, consciousness, and supervenience. Can we say that there is something that holds all of them together with some methodological bind? Dretske's answer is that we can. Introspective knowledge is a displaced perception (44); absent qualia has no relation, and hence they do not contextualize, but they are experienced as such.

When S represents F of K, it implies a representational fact. That is for some F, S represents the F of K. But it also implies something that is not a representational fact (possibly misrepresents), and what is not a representational fact is that which does not stand in relation C to S, and so they are hybrid of representational facts and facts about representation (26). External observers are often better positioned to observe a system's internal representation than is the system itself (48). This much is taken to be obvious. The first person authority can thus be jeopardized (54), and Dretske defends such a position against criticisms (Chapter 5), taking it more like a safety valve than an option. The Representational Thesis is assumed to be an externalist theory of mind (124).

For absent qualia, one needs to distinguish between epistemic use (Cf Jackson) and its phenomenal use. The epistemic use tells us that a dog that looks like a poodle to S, is what S believes to perceive, in the absence of countervailing circumstances. The sense is not strictly called epistemic, but epistemic in the 'doxastic' sense (67), and hence its corresponding phenomenal sense will include an additional 'discriminatory' clause, which says that it looks different to S from other objects of a similar type (68). This is just to prevent the people who are colour blind from saying of an object of another colour that it looks like a red object. This clause is deliberately left vague just to allow context-sensitivity and circumstantial relativity (68). Thus we can identify qualia with phenomenal qualities, that is, the way things appear to be, but with a representation clause (s). The representational device is not the same as it would have been if it had been in the vertical state. To some extent this agrees with the view that qualia is not functionally definable (Ned Block and Fodor, 1972; Sydney Shoemaker, 1991), yet it does not mean that they are physically definable. They are physically



definable as long as there is a description in physical terms, of the conditions in which systems have information-carrying function. What is captured by the naturalistic theory of indicator functions is sufficient enough for a naturalistic theory of representation or misrepresentation. For they become apparent in the way two representational devices are related to the external objects. In one, there is an indicator function *C* and in the other it is not.

Dretske's representational theory of consciousness builds on the critique what he calls Higher Order theories. The major shortcoming of Higher Order theories is that they differentiate between lower and higher orders of the mind (we are reminded of the modularity view advanced by Fodor in this connection). On Dretske's understanding a distinction such as the above is only a form of *reductio* because it assumes the relation and then proceed to prove that to be differentiable. They either fall a prey to epiphenomenalism, (117), or they cannot explain consciousness. Dretske's view seriously considers experience and thought as forms of representaton that are located horizontally rather than vertically and hence he is opposed to both the Higher Order theories of experience (Armstrong 1968, 1980) or Higher Order theories of thought, (Rosenthal 1986, 1990, 1991, 1993). the main casualty here is the definition of consciousness as 'consciousness of', and Dretske changes this into 'conscious with' to his own advantage. Supposing we say, we see *s* is just equivalent to saying that we are conscious of seeing *s*, which gives us (i) we are aware of the object we see (conscious of); and (ii) we are conscious of this awareness (conscious of conscious of seeing *s*). This is 'peculiar' (100).

'They are states (Dretske has a point against Rosenthal's distinction between creature consciousness and state consciousness) that make us conscious, not states that we make conscious by being conscious of them' (100).

Consequently, all experiences do not make us conscious of some object (hallucinations do not); but they are also veridical in a sense; so it follows that we are 'tethered' to an object all the time. Sometimes we have no contextual relation; so it is possible that our intentional representation says (or means) that we see pink rats but there is no outside object (pink rats).

Dretske defends an externalism that stems from the writings of Tyler



Burge. On Burge's view, he does not deny internalist positions. The question whether he reconciles internalist and externalist positions does not arise in the context as Dretske is against any such reconciliation (such a reconciliation is attempted by Akeel Bilgrami, 1992). Consequent upon the above considerations, Dretske's externalism bears a division into externalism about experience and externalism about thought. His defence is, therefore, against thinkers who agree with an externalist theory of thought but cannot reconcile themselves with an externalist theory of experience. Consistency demands that 'if an externalist theory of thought can be true, an externalist theory of experience can also be true' (127). This is a modal argument about externalism. The relation is not one in which implication between the antecedent and consequent holds. But, if a conceptual externalist (externalist about thought or belief) can accept, then he is also obliged to accept the point about phenomenal externalism (externalist about experience). Since phenomenal experience must be grounded completely externally, it is true one may not be aware of one's own qualia. Does it mean one should Quinise qualia in Quine's way so as to extract a theory of other ascriptions from the above? (Kanthamani, 1998). Self knowledge does not come from the privileged information one has but from extrinsic factors. That one does not look inward to get it (149). So, from this, Drestke was led to an argument about replacement for absent qualia, and supports it with a point about evolution without committing himself to *ceteris paribus* clause on the one hand and supervenience on the other. The causal pattern of law is not to be given in the *ceteris paribus* clause, but

Natural Selection → [Warm days → Chemical Activity → Colour Change].

For the case at hand, natural selection causes one thing which causes yet another.

Though this is not a solution to the nature of *ceteris paribus* laws, nor about supervenience, it is presented as one of the most plausible forms of naturalism. Thus, there is a counter to the earlier entailments which advocate a combination of supervenience and tractably sufficient as the hallmark (e.g., Baker), which advocates an entailment that combines supervenience with not tractably specifiable conditions (please note that it does not deny that they are

computed, or tractably specified). It combines tractably specifiable with not tractably specifiable, and brings back to our memory the Kripkean constellation, if not a Kripkean paradigm.

A. KANTHMANI

### REFERENCES

1. Andy Clark and Josefa Toribio, 'Doing without Representing?' in *Synthese* 101 (1994) pp. 401-431.
2. R. Cummins *Meaning and Mental Representation* (Cambridge : Bradford, 1989).
3. M. Devitt 'A Narrow Representational Theory of the Mind' in *Mind and Cognition*, (ed., W. Lycan, Oxford, Blackwell, 1990) pp. 371-398.
4. F. Dretske *Naturalizing the Mind* (Cambridge : Bradford, 1995).
5. S. Eynine *Donald Davidson* (Polity Press, 1991).
6. J. Fodor *Psychosemantics : The Problem of Meaning in the Philosophy of Mind* (Cambridge : Bradford, 1987).
7. *Theory of Content and Other Essays* (Cambridge : Bradford, 1990).
8. T. Horgan 'From Cognitive Science to Folk Psychology : Computation, Mental Representation and Belief in *Philosophy and Phenomenological Research* (1992) pp. 449-484.
9. A. Kanthamani 'Naturalizing Misrepresentation' in *Philosophical Quarterly* (1998) pp. 31-40.
10. S. Shoemaker 'Qualia and Consciousness' in *Mind* 100 pp. 507- 524.
11. S. Stich *From Folk Psychology to Cognitive Science : The Case against Belief* (Cambridge : MIT Press, 1983).
12. ----- 'What is a Theory of Representation?' in *Mind* (1992) pp. 243-261.
13. K. Strenelny *The Representational Theory of Mind* (Oxford : Blackwell, 1990).  
(Dretske 1995 contains all the other references)

## *Appeal*

The Indian Council of Philosophical Research has been planning to publish a Subject and Author Index of all the Philosophy Journals in the English language which have been published in India during the last several decades. The work has been entrusted to the eminent librarian, Dr. S. C. Biswas, who has been trying to collect the relevant copies of the Journals from various individuals and libraries all over India. However, inspite of his best efforts, the following copies of the Journals have still to be located so that the work, which is an important reference tool, could be completed.

We appeal to all individuals/institutions to find if they possess these missing issues in their own collection and in case they have, content pages of volumes/issues, they may please get the same photocopied and send them along with the bill of expenses to the Indian Council of Philosophical Research, Rajendra Bhawan, Fourth Floor, 210 Deen Dayal Upadhyaya Marg, New Delhi 110002. Their help in this regard will be specifically acknowledged and a copy of the comprehensive Subject and Author Index, when published, will be sent to them as an expression of the Council's gratitude.

### ICPR - CUMULATIVE INDEX OF PHILOSOPHICAL JOURNALS PROJECT List of Indian Philosophical Journals

*Darshan Samachar/Philosophica*

Vol. 7 (4) to issues upto 1991

*University of Allahabad Studies (Philosophy section)*

1940, 1944, 1945, 1946, 1947, 1949, 1953, 1959, 1960, 1961, 1962, 1969.

*Research Journal of Philosophy*

Vols. I (1), 2 (2), 3(2), 4 - 8(2), 9(1978) - 1991.

*Anviksika*

Vol. 7 to issues upto 1991.

*Indian Philosophical Annual*

Vols. from 1989 to 1991.

*Indian Philosophical Research*

Vol 3 (1920) to issues upto 1991

*Indian Journal of Philosophy*

Vols. 5(1964) to issues upto 10

*Indian Philosophy and Culture*

Vol. 21(1976) - 1991

*Indian Philosophical Congress*

1960. 1962-1964. 1969-1973, 1975-77, 1980-1983, 1985, 1988, 1990, 1991.

*Journal of the Philosophical Association.*

(Incorporated Indian Journal of Philosophy) Vol. XI. no. 40(Dec. 1968).

Vol. 15(1974) to 1991.

EDITOR