

# THE EXTERNAL MIND

---

An Essay Concerning the Relation between Psychological  
Content and the Nature of Things.

A Thesis For the Degree of Master of Philosophy

~~EXTERNALISM~~: An Essay on the Nature of  
the Relation between psychological  
States and the World.

by

M. R. J. Jennings.

Department of Philosophy

University College,

University of London.

April 30th 1995

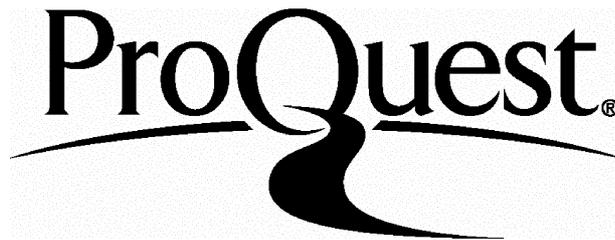
ProQuest Number: 10018655

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10018655

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.  
Microform Edition © ProQuest LLC.

ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

## Table of Contents

---

Abstract .....	1-3
Introduction .....	4-8
Chapter 1 .....	9-19
Chapter 2 .....	20-30
Chapter 3 .....	31-45
Chapter 4 .....	46-59
Chapter 5 .....	60-69
Chapter 6 .....	70-83
Conclusion .....	84-85
Bibliography .....	86-88

## Abstract

---

This thesis examines the assumptions behind various of the individuation principles of propositional attitudes and other Intentional states - *i.e.* those principles by means of which the sameness and difference of psychological states are determined. It argues against the Externalist doctrine that the content of our mental lives depends for their individuation on states of affairs in the world: that is to say, on those things in the world one's thoughts or experiences are purportedly about. How the world happens to be matters little, I suggest, to the kind of thoughts and experiences one is capable of having.

In doing so I provide a (I hope useful) critical survey of some of the more influential writings available on the subject. The thesis begins with a summary of Putnam's main thoughts concerning the semantics of natural kind terms, and examines the development of Burge's Externalism from his singular thought theory (1977) to that concerning general thoughts.

Nowhere, however, do I align myself whole heartedly to any one philosopher. I reject, for example, the so-called "middle way" between Internalism and Externalism, advocated by Fodor and others: that view which accepts that truth-conditions are externally individuated but which argues that a new theoretical construct (narrow content) be devised in order to accommodate the otherwise plausible intuition that mental states be individuated by their causal role. The proposal is criticized for its obscurity.

Rather than explore other versions of the Dual Component hypothesis, I propose instead that we reject Externalism *tout court*. The content of one's concepts are not individuated by one's social allegiance, or by the nature of things in the world which they subsume. No: one's concepts are often one's own constructs. Furthermore, I show that we can plausibly accept the consequence of immunity from conceptual error. I then summarize the debate between G. Segal and M. Davies, showing how the private concept strategy can answer Burge's non-linguistic Thought Experiment.

Lastly, the Russellian Thought theory of Gareth Evans is looked at briefly, as are various responses. I argue that singular thoughts are not *sui generis*, but can be collapsed into general contents, and that their truth-conditions can be given partly in terms of our experiences rather than entirely by those things in the world which we assume to be their cause. I then defend Internalism about general thoughts by drawing attention to the implausibility of a consequence, otherwise neglected or ignored, of Burgean property-dependent Externalism: namely that such thoughts would not be available in wide scale hallucinatory scenarios.

## Introduction

---

Let us say that a person has an Intentional state whenever that person thinks, believes, wants, hopes, fears (etc.) or experiences something, whatever that "something" might be, and whether or not that "something" exists. Then Intentional states are *ipso facto* always directed at, are always about, something, such that, insofar as it has this "directedness", it can be ascribed a "content" in terms of which its sameness and difference from other Intentional states can be discerned.

We can liken the notion of mental content with the notion of a thought or proposition, (say) of the form *a is F*, whose truth-value (its truth or falsity) will be determined by such conditions as, for example, that *a* falls under the concept *F* (or according as *a* has the property *F* predicated of it). Now insofar as any thought or proposition shares the same truth-conditions the thought or proposition will be the same. Further, let us say that the act of believing something involves taking an attitude towards a proposition, that its content is either true or false, and that an Intentional state can be characterized in terms of this notion of a propositional attitude.

Some Intentional states, however, are less clearly characterizable in terms of their propositional form. Here we might instead take mental content as being *whatever it is that our experience or mental state is about*; that it is the "what it is like" to have an experience. In this way we can characterize the state by its "phenomenological" properties. In general, however, we can say that an Intentional state is quintessentially representational; it represents something and does so in some way.

Now there is a debate among philosophers concerning just how the individuation of Intentional mental states might proceed. By *individuation* I mean that principle by means of which we ascertain the conditions of identity and existence of an entity; in our case a mental content. So when I ask after the individuation conditions of something I'm asking (a) what is essential to the sameness or difference of that entity - *i.e.* what would it take to change its nature or identity - and thus (b) what makes the existence of that thing possible. Clearly, the principle is fundamental to any investigation of any

object or event. For one must be clear about what "constitutes" the state or thing under investigation. And knowledge of what constitutes the entity in question will be available when, but only when, we know what that entity depends upon for its individuation. The trouble is that there are several such principles of individuation: we will examine each in turn.

One reason why we might be concerned with knowing the identity conditions of mental content is because it facilitates the use with which we put the notion of Intentional state: namely as a theoretic construct in the explanation of action. To explain a person's action is to explicate its cause, and this, we commonly suppose, consists of a reason, itself analyzable in terms of a conjunction of beliefs and desires (*i.e.* propositional attitudes). So for example, that Alf swam the distance was *because* (a) he believed that the swim would help keep me fit, and (b) *because* he wished to stay strong.

Another reason for our interest is metaphysical. Here the hope is that knowledge of the individuation conditions of an Intentional state might afford us knowledge of the nature of the relation between the mind and the world. Some philosophers believe that the relation is essential to the nature of our experience and thought; others deny this. So for example, we have on the one hand the Externalist doctrine that mental content is individuation dependent on the subject's environment in a "deep and fundamental" way. What this relation means for the Externalist is, amongst other things, that where there is a difference in the object or property in the environment there is a corresponding difference in the subject's Intentional states. The idea here is that the nature of these external events and states of affairs help to constitute the content of the mind. By corollary, since content individuates mental states, the *nature* of mental states too are constituted in terms of the things external to the subject's body. Otherwise put, if we assume that the content of a mental state determines the truth-conditions of the propositional attitude in question - in other words, that the contents themselves specify what facts must obtain if those contents are to be true - then if these truth-conditions differ according as the environmental states of affairs differ, then so too will the mental contents. In order then to capture the Intentionality of a mental state,

the Externalist urges that we include as an essential part of what it is about that very object or property in the world.

Of course to say that some mental states are constituted by external states of affairs is not to say that individuation-dependence is such that the mental state is literally to be identified as some external physical fact. No: the intuition is that my mental states can be about, e.g., rocks and trees, etc., only if there are such things in my environment to fix the identity and existence conditions of those states. As we will see below, it is said that a person must be in, or have been in, or must rely on a group or member of his linguistic community that has been in, some contextual relation with the relevant state or thing in the environment. But of course the rocks and trees (etc.) are not themselves the mental states: they simply contribute towards the latter's Intentionality.

Internalism, on the other hand, is the denial that mental content is individuated in anything like the above terms. Not for the Internalist do external facts enter as constitutive of one's mental life. There is no deep or fundamental individuating relation between (a) the kind of mental states an individual can have and (b) the nature of the things in his environment. Rather content is individuated in terms of the subject's internal states; these might include one's behavioral dispositions, proximate stimuli or phenomenology, one's acts and brain states. Only such facts as these constitute mental content.

The Internalist makes what philosophers call a strong local supervenience claim: to wit, there is no difference in one's Intentional states, given one's chemical and neural or functional history, without a difference in one's physical states. Vindication of Externalism therefore rests on a refutation of this strong local supervenience claim. The Externalist must show that for physical duplicates  $x$  and  $y$ , embedded in their respective (and qualitatively different) environments  $w_1$  and  $w_2$ , it is possible that they could differ in their Intentional properties.

Chapter 1 sets the scene by expounding Putnam's Externalism concerning the Intentionality of natural kind terms: to wit, that their reference be explicated in terms of a causal link with the essences of the things referred

to, a thesis which urges against the Descriptivist or Internalist view that reference is achieved where something falls under the concept which the subject associates with the name. That chapter continues by showing how Burge extends Putnam's strategy to the thesis that not just the content of words but the content of our Intentional states are affected by one's environment, specifically one's social affiliations. Chapter 2 looks more closely at Burge, and outlines his other strategies; in particular, it traces and clarifies the development of his Externalism, and in doing so highlights some confusions that this involved.

Having expounded and clarified the Externalism of Putnam and Burge, chapter 3 examines one kind of critical response: namely Fodor's version of the Dual Component Theory. This kind of theory accepts that truth-conditions are individuated by the environment (*i.e.* it accepts Externalism), but holds that our notion of content can be taxonomized into two kinds: that it is "broad" when individuated externally, and "narrow" when individuated internally. The dual component theorist maintains that since it is "narrow" rather than "broad" content that is important for psychological explanation, the notion of external individuation of mental states is largely without significance. My examination shows that Fodor's proposal concerning "narrow content" is obscure.

That obscurity results, as Chapter 4 points out, from accepting part of the Externalist moral. Here I consider a different notion of narrow content, that put forward by T. Crane. He urges that we reject Externalism *tout court*: truth-conditions are individuated internally, not by the world. This view involves possession and formulation of private concepts; I defend this idea largely by default: for the criticisms against it are not as strong as they might at first sight appear. The rest of that chapter shows how a similar strategy, put forward by Segal, answers Burge's non-linguistic Thought Experiment.

Lastly, I look at another version of Externalism: Gareth Evan's Russellianism: that view which argues for a species of singular thought whose existence and identity is tied to an external object. Suppose I see a lion before me, whereas my doppelganger hallucinates one, that we both utter "I want to capture it", and proceed to act thereby. Russellians are committed to the view

that whereas I have a singular thought expressed by that sentence, my doppelganger does not. Against this, I consider the view that Russellianism is unmotivated: that because the behaviour and experience of mine and my doppelganger's appears to be the same, we have every reason to suppose that the thoughts are too. I conclude that the arguments presented in this vain do not decide the issue.

Moreover, I offer reasons to doubt the alternative view that singular thoughts, though not Russellian, are nevertheless irreducible to general thoughts. My doubt arises when we try to individuate the thought had in the hallucinatory case. Instead I defend the view put forward by Blackburn and Searle, that all thoughts are essentially general, and in particular I defend (also by default) the latter's version of how thoughts had in hallucinatory scenarios can be individuated in terms of determinate truth-conditions. Finally, I answer the retort that though Searle's (and Blackburn's) Internalism subverts Russellianism, it begs the question against Burge's Externalism concerning general thoughts. My response is to ask the Burgean to account for the kind of thought had in a hallucinatory scenario: that is to say, I draw attention to the *prima facie* implausibility of general thought Externalism, by suggesting that their putative property-dependence makes them susceptible to illusory thoughts.

Thus I defend by default what most philosophers in the debate would regard an extreme form of Internalism. Short of a knock-down argument - rather than reliance on intuitions (which is everywhere apparent in the forms of Externalism examined below) - I see no good reason to suppose that the nature of the relation between the mind and the world is deep: no reason to suppose that it bears on the kinds of thoughts and experiences we can have: no reason then to reject the view that all our thoughts and experiences could be had however the world might turn out to be.

## Chapter One

---

**§1. Fregean Sense and the Puzzle of Identity.** Consider the mental life of Oedipus. He longed to marry Jocasta, only Jocasta was his mother (little did he know). There is a sense in which we can attribute contradictory thoughts to him: that he wanted to marry Jocasta but not his mother. Yet we wouldn't say that he was irrational. Rather we would say that what accounts for his not knowing the identity involved is that he looked upon the same person from different perspectives.

It is just this phenomenon of perspectivity which gives rise to that puzzle of Frege's which asks how a statement of the form (i) [a = a] can be uninformative (uninformative because it is an analytic *a priori* truth: a statement whose truth is known independently of experience and in virtue of the meaning of the terms involved) whereas a statement of the form (ii) [a = b] can represent a significant addition to our empirical knowledge (and is thus synthetic *a posteriori*), even although "a" and "b" denote the same object in the world?<sup>1</sup>

The resolution of that puzzle comes, we are told, with the introduction of Fregean *senses*: *i.e.* the constituents of content, a thought or proposition. Now Frege is often attributed the view that the *sense* of a proper name is, *inter alia*, a description (where the description accounts for our conception) associated with the object or bearer of the singular term. The above puzzle of identity can then be accounted for in terms of the different *senses* (different information values) associated with the proper names: these proper-names stand on either side of the identity statement, making it informative thereby. Here the criterion for sameness or difference of sense is determined by whether a thinker can without irrationality take different epistemic attitudes towards the identity statements: *i.e.* whether the thinker can reasonably hold conflicting beliefs about, say, the statements (i) and (ii).

Explanation of identity statements is not, to be sure, the only function the notion of sense is deployed by Frege to serve. And indeed it is this further

---

<sup>1</sup> Frege himself believed that analytic *a priori* truths could be informative; Kant, on the other hand, believed that only synthetic *a priori* truths could be.

function of sense which, as we will see, most exercises Externalists and Internalists in their theories of reference.

**§1.1. Semantic Internalism: Descriptivism.** To the question, in virtue of what does a name (or other term) refer to a concrete object, there are, broadly speaking, two kinds of answer one might give: a Causal theory, or a Descriptivist theory.

Descriptivist theories explain reference in terms of the following:

- (A) *there is a relation between the expression and a purely general concept<sup>2</sup> or a cluster of concepts, and this relation can be grasped by the mind;*
- (B) *that the relation between a purely general concept and referent is a relation of fit.*

Here the speaker's ability to refer is explained internally - as Putnam would say, by what's in the head - such that success in using (say) a proper name lies in one's ability to associate it with the right descriptive concept, together with there being an object to satisfy it.

Putnam holds, however, that this semantic theory leaves out two apparently important components: other people and the world.<sup>3</sup> Moreover, he holds that assumptions (A) and (B) can be recast in a way that reveals their incompatibility.

Note, first, that the above assumptions correspond to the following two semantic concepts:

- (i) *Meaning conceived as Intension;*
- (ii) *Meaning conceived as Extension.*

On the relation of fit, we take it that the extension of the term "rabbit" is the set of things (rabbits) the term is true of. We take it also that two terms can have the same extension but differ in meaning: e.g., the compound terms "creature with a heart" and "creature with a kidney" (assuming that creatures

---

<sup>2</sup> Sense, intension, non-Fregean concepts (and meaning: see below) are used interchangeably throughout.

<sup>3</sup> H. Putnam, "The Meaning of Meaning" in his Philosophical Papers, *Mind, Language and Reality*, Vol 2 (1975). p. 271.

with one have the other). So by the concept or description associated with the term we mean intension.

With this in mind, Putnam reformulates assumptions (A) and (B), thus:

(i) *knowing the meaning of a term is a matter of being in a certain psychological state*;<sup>4</sup>

By "psychological state" Putnam means amongst other things the kind of state involved in (say) knowing how to factor numbers, states of memory and psychological dispositions. So *knowing* the concept is a psychological state of *understanding* the intension.

(ii) *"the meaning of a term (in the sense of intension) determines its extension (in the sense that the sameness of intension entails the sameness of extension)"*.<sup>5</sup>

Given this reformulation, it follows that psychological states must also determine extension. For if the meaning of a term determines its extension, and if knowing the meaning of a term is determined by being in a certain psychological state, then psychological states must also determine the extension.

**§1.2. Semantic Externalism: A Causal Theory.** Putnam offers three counter-examples - we'll examine each in turn - to the effect that no notion of meaning can jointly satisfy the above two assumptions of Descriptivist semantics. For according to Putnam, the extension of a term can be different in the idiolects of two subjects, even although psychologically the subjects are the same: this because the meaning of a natural kind term is determined not by the subject's psychological state, but by the "essence" of the thing referred to (see below).

Putnam was not alone in his attack on traditional semantics. Kripke, Donnellan and others, all made broadly similar claims, and some propounded an alternative causal theory of reference. The core idea of the theory is that, contrary to traditional semantics, conceptual competence alone is insufficient

---

<sup>4</sup> Op. cit. p. 221.

<sup>5</sup> Op cit. p. 219.

to explain the speaker's ability to refer; that there must be an external *causal relation* between speaker/s and referent - even although that relation might elude the speaker's epistemic reach. In other words, it matters not that one has never stood in a relation of acquaintance with the referent or extension itself, for one might have learned the name or predicate from others who have stood in the appropriate relation; or if not them then others from whom they acquired the name, or so on down the causal line of name users. Now in the spirit of this causal theory of reference, Putnam proposed to reject (i), and to revise (ii) so that, though meaning determines extension (assumption (ii)) meaning itself is not determined by psychological states (assumption (i)), but by the nature of things in the environment.

Here then are Putnam's counterexamples.

*Example 1.* Imagine a world identical in every respect to ours (including its inhabitants) except that its water, though it seems phenomenologically the same as ours (it looks, tastes, feels and functions like ours), consists not of H<sub>2</sub>O but of a different microstructure, XYZ. When in our respective worlds my doppelganger and I point to a lake and utter "that's water" we make the same type of sound, use the same type of sentence, but mean, according to Putnam, quite different things. For our terms have different extensions, they pick out different things: Twin Earth's lake is XYZ, not water (H<sub>2</sub>O). Yet *ex hypothesi* our psychological states - our intensions and the concepts associated with them - are the same. So psychological states do not determine extension.

A further word about "psychological states" before considering the other counter-examples. Some philosophers of semantics have assumed what Putnam calls "methodological solipsism".<sup>6</sup> Solipsistic states presuppose the existence of nothing beyond the subject (not even one's body). Putnam has dubbed these states "narrow" in contrast with "wide" states, which do carry existential commitment. He includes in his understanding of narrow states such things as hallucinations, sense-data, psychological dispositions, visual

---

<sup>6</sup> Op. cit. p. 219-221.

experiences or phenomenological properties. Partly for this reason, the concept of intension, like concepts generally, has been regarded by traditional semantics as something like a mental entity. It is because of this assumption about "narrow states", Putnam points out, that traditional semantics leaves out the two [allegedly] important things: other people and the world (see below).

Putnam's main reason for bringing out the assumption of methodological solipsism is that it is much more difficult to show how "narrow" states can plausibly be said to determine extension than it is for "wide" states. For "narrow" states, because *ex hypothesi* they are fixed by the intrinsic properties of the individual, are not had in virtue of their relation to any external environmental cause or effect. With "wide" states, on the other hand, my intention to use "water" may be to refer to water, whereas my twin's intention to use "water" may be to refer to twin water. Our intentions differ in having different Intentional objects. Hence it would not be implausible to say that our intention determines the extension. So Putnam's thesis that meaning is not determined by psychological states rests on holding psychological states constant: he can't allow that XYZ or H<sub>2</sub>O affect them.

*Example 2.* Suppose that only experts can distinguish between molybdenum and aluminum; that molybdenum, but not aluminum, is common on Twin Earth; that aluminum, but not molybdenum, is common on Earth. Now whereas on Earth pots are made of aluminum, on Twin Earth they're made of molybdenum. Suppose we switch names on Twin Earth so that "aluminum" refers to molybdenum, and "molybdenum" refers to aluminum. If an average speaker goes to Twin-Earth he would take it that "aluminum" pots were made of aluminum. As indeed Twin-Earthians would vouch. The difference, of course, would not elude the Earthian metallurgist. Similarly, the Twin-Earthian metallurgist would know that aluminum was "molybdenum". Although on using the word "aluminum" the average Earthian and Twin-Earthian speakers would have the same psychological states, the extensions of their words would be different: they would mean different things.

*Example 3.* Suppose you and I are unable to tell a beech tree from an elm; that you use "elm" and I use "beech". The two terms nonetheless have

different extensions, for "elm" applies to all elm trees, "beech" to all beech trees. The idea is that this inability to discriminate implies that our concepts are identical: they will be something like *big deciduous tree*. Hence the difference in extension is not borne out by any difference in what's in our heads. Now suppose you are my doppelganger in a counterfactual world such that our thoughts, dispositions, and the phenomenology underlying those dispositions, are indistinguishable. Further, let us switch names again. In the counterfactual world "elm" refers to beech, and "beech" refers to elm. Surely, we mean different things. Surely, when you say "elm" you mean beech; and when I say "elm" I mean elm? Yet *ex hypothesi* we are psychologically the same.

**§1.3. Other People and the World.** The general point then is that meanings aren't, as Putnam puts it, "in the head". Putnam's causal theory of reference has it that the meanings (in this case the state of knowing the meanings) are individuated in terms of the historical link between oneself and the referent of the term, rather than, as the semantic Internalist would make out, solely in terms of one's conception of things. Still, it is odd to say that meaning isn't "in the head". In fact Putnam needn't say this. For suppose that both your arms have burn marks, one from the sun, the other from the stove. However similar they look, their causal origins makes them distinct. Yet both states are in the arm, both are intrinsic states of yours.<sup>7</sup>

Putnam's last two counter-examples make an additional point. In the first example, the contribution of the environment - Putnam's essentialism - is borne out by the fact that the substances were as different in the two environments before the dawn (in both worlds in 1750) of modern chemistry as they are now - that the denotation of "water" before this scientific revolution was H<sub>2</sub>O, just as it is now - and that the discovery of the nature of the substances is therefore just as much a discovery of the meaning of the

---

<sup>7</sup> Cf. Davidson: "On Knowing One's Own Mind", *Proceedings and Addresses of the American Philosophical Association* (1987). Stalnaker makes a similar point in "On What's in the Head", in *The Nature of Mind*, ed, D. Rosenthal.

terms. It's not the case, in other words, that the difference in the meaning of "water" between Earth and Twin Earth depends on the fact that some people on both planets can tell the difference between the two substances.

In examples 2 and 3, on the other hand, our acquiring the names is more explicitly a function of the fact that there are members of our language community who, unlike you and I, have a *method of recognizing* whether the referents have the features that help indicate membership of the relevant extension. This phenomenon - Putnam calls it the "division of linguistic labour"<sup>8</sup> - accounts for why, in the last two examples at least, the two assumptions of traditional semantics are not jointly satisfied. For when the average speaker (he who defers to the expert for want of a criterion) acquires a word, it is not his psychological state that fixes the extension of his term - again, myself and my doppelganger are psychologically the same - but the experts, for it is they who have the discriminatory nous to say whether a concrete object falls within the extension.

Of course their *method of recognizing* needn't consist in recognition of the microstructure of things: "experts" about water were not around before the development of chemistry in the 18th century; they need only be able to distinguish these markers and set of descriptions.

But what is to say that "water" has the same extension and use before and after this scientific revolution?

What happens is this. The average speaker typically acquires a term for a natural kind by being shown "markers", and perhaps also by being given a set of descriptions common or stereotypical of the referent. Now suppose I'm introduced to the term "water" and that, pointing to a "marker", the expert gives the ostensive definition, "this is called water". The ostensive definition, says Putnam, involves an empirical presupposition.<sup>9</sup> The presupposition is that the stuff identified by the "marker", the stuff pointed to, bears a similarity relation (*same<sub>r</sub>*) to the stuff which the other members of the linguistic

---

<sup>8</sup> Op. cit. p. 227.

<sup>9</sup> Op. cit. p. 230.

community call "water". The *same<sub>L</sub>* indicates a necessary and sufficient condition for membership of the extension of "water": it stipulates that the sample must share the same microstructure as the previously identified samples. Of course the sample might turn out to have a different microstructure: it might be gin or XYZ. Then the ostensive definition fails. Hence it is defeasible; it rests on the empirical presupposition that the *same<sub>L</sub>* condition has been met. Thus before the scientific revolution, the samples of liquid were distinguished by their superficial qualities - their nominal essence - and assumed to share a common real essence. And though the subsequent scientific investigations have confirmed that they do share a common microstructure, the ostensive definition is still defeasible: future investigation may force a revision.

In saying that "water" is the stuff that bears a certain similarity relation to the water around here, Putnam is claiming that "water" and indeed all natural kind terms have an unnoticed indexical component. He is exploiting the common belief that the two assumptions of traditional semantics elude application to ordinary indexicals or token-reflexives. For generally speaking, the intension and associated concept of ordinary indexical and token-reflexive terms remain constant while the extension varies from context to context. So for example, when I and my counterpart utter "I'm hungry" our concept associated with "I" is the same. Yet the extension of "I", being a function of whoever utters it, is different in the two cases. Putnam thus generalizes the point to natural kind terms (*i.e.* where my twin and I have indistinguishable concepts) such that in our respective worlds "water" is the stuff that bears a certain similarity relation to the water around here. But the similarity relation is different: so then is the extension.

Otherwise put, the natural kind term, defined by the *same<sub>L</sub>*, is a rigid designator: so defined, our use of "water" means the same in all logically possible worlds in which water exists: a substance in world  $W_2$  which shares the same microstructure to the substance in  $W_1$  - has the *same<sub>L</sub>* - is therefore water. So "water" refers to  $x$  if and only if  $x$  is the same (liquid) as *this* (where "this" indicates that originally identified stuff which other samples must

match in microstructure). As Kripke says, such statements, being true in all possible worlds, are metaphysically necessary: water can't be anything other than what it is; and what it is has been discovered to be H<sub>2</sub>O.

What of the contrary intuition that water could turn out to be something other than H<sub>2</sub>O? According to Kripke, the intuition conflates metaphysical with epistemically necessary truths.<sup>10</sup> We can say truly that it is epistemically possible that a substance with the same phenomenological properties as Earth water could turn out to have an essence or micro-structure other than H<sub>2</sub>O. That would explain our intuition that water ≠ H<sub>2</sub>O; but it wouldn't be water: water = H<sub>2</sub>O. The conflation arises, we're told, because a description of the phenomenological properties of water could have been used in fixing the reference of the name, and moreover that this description is often used to identify the referent or extension in question. It is this then that lends itself to the idea that the reference fixing description is *synonymous* with "water", and with it the idea that the substance has these properties only contingently. Nevertheless, it is apparently a necessary fact that water has the essence that it does. So the meaning of "water" isn't synonymous with the reference fixing description. After all XYZ satisfies the operational definition, but fails to bear the *same<sub>L</sub>* relation. Rather the meaning of "water" is the referent.

It is important to note that given Putnam's indexical thesis, he could have denied that intension determines extension, rather than say, as he did, that a difference in extension is a difference in intension, and thus that meanings aren't concepts (mental entities) fixed by the psychological states of those who use them: my concept and my counterpart's concept is in a sense the same, for we can't distinguish between the two substance with the same nominal essence. But the intension with which the words are associated are different, since - and here he retains the idea that intension determines extension - they refer to different things. Hence, unless the referents are the same, the intension cannot be. Therefore our words mean different things. But then it will be wondered what the intension of "water" is if it is not the description

---

<sup>10</sup> 1972, p. 157.

or concept my twin and I possess. Of this see more later (chapter 4).

**§1.4. Individuation and Intentional Properties.** What goes for semantic properties goes for Intentional properties, according to Burge. Deploying the same structure of argument as Putnam, Burge attempts to show in his "Individualism and the Mental"<sup>11</sup> that just as semantic properties are fixed by the subject's environment so are Intentional properties or mental content, though only one of Burge's strategies - he has three of them, each of which will be examined in turn - relies on the idea, adumbrated by Putnam, that meaning is fixed by the socio-linguistic community. We shall see, however, that for all the similarities between Putnam and Burge there are significant differences too. In the meantime, let me finish this chapter by showing how Burge extends Putnam's argument.

(a) Of all Alf's beliefs about arthritis all but one is true: he thinks [falsely] that he has arthritis in his thigh. He reports to his doctor; his doctor corrects him, tells him that arthritis is an inflammation only of the joints.

(b) In a counterfactual situation, Alf<sub>CF</sub> is identical to Alf in his non-Intentional mental and physical history; only "arthritis" in *his* social environment means an inflammation of the thigh as well as the joints.

(c) In neither situation then are their concepts the same. So in neither situation can their Intentional content be. Yet *ex hypothesi* they are internally - constitutionally - the same. Hence: where but in their social affiliations - what with the experts' "individuating" or "distinguishing" knowledge *vis-a-vis* the communal concepts - can the difference between them lie?

Clearly, this strategy, if sound, applies, just as Burge maintains, to the use of all terms, not just to natural kinds.

All that matters - and this is the essence of Burge's strategy - is that it be possible for the subject to be in systematic error in the use of his

---

<sup>11</sup> See T. Burge, "Individualism and the Mental", *Midwest Studies in Philosophy* vol. IV, pp. 73-121, 1979.

concepts. Remember: error lies only with a constituent or notion<sup>12</sup> of the concept (until corrected by an expert, he continues to believe that this is the right application, *i.e.* that the concept "arthritis" applies to the ailment in his thigh), that it is this which makes for Alf's partial understanding of the concept *arthritis*. So this kind of error is possible because the subject has mastered the other notions (which make-up the concept) as his correct application in those other contexts shows. Thus the error isn't just a careless one-off misapplication; were that so it would be less plausible to maintain that the Alf is mistaken in what he believes.

Hence, against the strong local supervenience claim mentioned in the introduction, there apparently can be a difference in mental content of twins without a difference in their brain states.

---

<sup>12</sup> Ibid. p. 75.

Burge substitutes "notion" for Fregean senses in order to show ontological neutrality.

## Chapter 2

---

§2. **BURGEAN THOUGHT.** Before considering Burge's Social Externalism further, I want to digress a little by outlining something of the evolution of his thought, as this will enable us better to understand not just the flavour of Burgean Externalism and some possible confusions that have arisen in its development, but also to understand some of the contours of the debate in general about individuation of content. In particular, I want to consider the discussion in his 1977 paper "Belief De Re"<sup>1</sup> of the distinction between two kinds of belief ascription, *de re* and *de dicto*, and to suggest that the kind of singular thought Externalism believed by Burge at the time to follow from this distinction led him later to reconsider his views and subsequently to advocate in "Individualism and the Mental" the much stronger thesis about the individuation of mental content, namely his Social Externalism.

§2.1. **DE RE AND DE DICTO: PROPOSITIONS.** As will be seen more fully when we discuss the singular thought Externalism of Gareth Evans, one's construal of *de re* propositions will depend largely on whether or not one takes Fregean sense to be adequate in determining the reference of singular terms, and, crucially, on how one takes Fregean senses themselves to be: in particular whether *sense* is tied to the object of one's thought.

Now in connection with *de re* propositions, Burge contends, *pace* Internalists, that *sense* is insufficient to determine reference, that a contextual, non-conceptual relation - namely to the object itself - must be specified in order to determine the *de re* thought. This is to say that since the truth-conditions are themselves determined by states of affairs, and since states of affairs are constituted by facts that obtain, the external environment is what individuates the singular thought. Before saying what this might amount to, let me say more about the *de re* and *de dicto* distinction, at least as Burge understands it.

Ever since Quine's<sup>2</sup> seminal discussion of *de re* and *de dicto*, the

---

<sup>1</sup> See T. Burge, "Belief De Re", *The Journal of Philosophy*, 1977.

<sup>2</sup> "Quantifiers and Propositional Attitudes", *The Journal of Philosophy* 1956.

common assumption has been that the distinction is best understood in terms of the "substitution criterion". The idea here is that it matters not at all on describing a subject's *de re* belief what description one gives of the object to which his belief is directed, so long as the description itself pertains to the thing believed. The reason is that the *de re* relation between object and the subject's belief is thought to be sufficiently direct as not to warrant ascription of any conception which the subject might use to pick out the object. By contrast, it is said that a *de dicto* belief ascription can fail of substitution of coextensional terms: *e.g.*, it cannot be inferred from x's belief that "2 squared is 4" to the belief that "2 is the only even primed squared of 4", since x may believe the first proposition, but not, because he fails to realize that 2 is the only even prime, the second proposition.

As it happens Burge shows that substitution can also fail with *de re* belief ascription: x may believe that the man at the corner is a spy, where the concept expression "man at the corner" relates x *de re* to the man referred to, such that if x's belief involves the denoting terms "the man at the corner", we should be unwilling to substitute any other phrase (*e.g.*, "son of the Kiev vodka distiller") true of that object. As Burge puts it, the concept-term "man at the corner" does double duty at the surface-level: it characterizes x's conception, and it picks out the relevant *res*.<sup>3</sup>

A better characterization of the distinction, he suggests, is the epistemic one: that unlike the *de dicto* attitudes, *de re* aren't fully conceptual because they require the external relation to the object for their individuation. That external relation will be contextual or non-conceptual as *de re* ascriptions involve expressions of indexicality: that is the truth-value of the propositions in which the expression occurs will vary with the context in which the expression is used. Burge illustrates this idea, that the conceptual element is insufficient to pick out the *res*, by having us imagine seeing a man coming from the distance in a swirling fog. Though we might be said to believe of the man that he is wearing a red cap it would be mistaken to say that the

---

<sup>3</sup> Ibid. p. 341.

description "wearing a red cap" - and hence the conceptual element of our belief - suffices to fully individuate him. The fog may be too thick, the distance too far, for our description or concept to pick out the object. And even if we could distinguish the man clearly, that is describe or conceptualize him fully, he might be indistinguishable from surrounding look alike, thus rendering the individuation of him only by a temporal-spatio, that is nonconceptual or contextual, specification.<sup>4</sup> (In another example, we imagine someone saying "I believe of the present moment that it is in the twentieth century". We have already seen that with indexicals the subject's conception may be insufficient to individuate what is being indexed: here the present moment.)

The distinction can also be given a semantic characterization. Here *de dicto* propositions are expressed in closed sentences; whereas *de re* propositions are expressed partly in open sentences and partly by what is expressed given the *res*. To say that *de dicto* belief ascriptions relate the subject to a fully conceptualized proposition is to say amongst other things that on ascribing a belief to a subject we need only make reference to the elements of his "conceptual repertoire": we need concern ourselves only with those concepts which make up the subject's belief content.<sup>5</sup> Alf believes that  $[\exists x (x \text{ is a spy})]$  (where the square brackets indicate a complete proposition). Conversely, we can express the idea that *de re* propositions are not fully conceptualized by saying that attributions of *de re* beliefs contain quantifiers outside the scope of the "belief" verb:  $\exists x (\text{Alf believes that } x \text{ is a spy})$ , such that *de re* thoughts involve a context (an individual or object) to which the subject must stand *en rapport* if the proposition in question is to be complete or fully interpreted. So for example, the subject position in the open sentence "x is black" is occupied by a free variable, such that only where the variable takes an object as a value (e.g. an individual cat) does the act of predication render the predicative fragment fully determinate or fully interpreted, that is,

---

<sup>4</sup> Cf. G. Evans. Also see chapter 4 below.

<sup>5</sup> Op. cit. p. 345.

can it be rendered fully propositional, as being either true or false, according as the indexically marked object in question has cat-hood attributes and has the further property predicated of it, namely blackness. For this reason Burge has called *de re* propositions "fundamentally predicational" (of this more later).<sup>6</sup>

It is important to note that Burge can still make use of *sense* as constituents of *de re* propositions, and so can still cut content finely: it is just that sense by itself is inadequate in determining the referent.<sup>7</sup> (The object in context is also a constituent of the proposition: only, to put it in Russell's terms, it is shown rather than expressed.) For the perspectivity of thought and experience occurs as much in the context of using demonstratives and ordinary proper names as in contexts of definite descriptions. I may know someone in two walks of life (e.g., by the name of John) without knowing they are the same person. (We can suppose that our only contact is by telephone, that when we speak *vis-a-vis* one walk of life the phone used is different: it is rather old and the reception of John's voice is different from other occasions, *i.e.*, conversations on another phone *vis-a-vis* the other walk of life.) I would find it informative to be told that they are the same John. Similarly "this = this" would be uninformative if the demonstratives referred to the same object under the same aspect or perspective, though I would find it interesting were the first demonstrative used with a gesture to (say) a picture of the Hope Diamond, and the second demonstrative with a gesture to a dirty stone.

We can begin to show the wider importance of this conceptual feature of *de re* propositions by considering in more detail the extent to which we have here an Externalism *vis-a-vis* singular thoughts. A central idea so far is that because *de re* propositions involve an indexical or contextual element, the description or concept contained in a proper name or singular term is thought insufficient to individuate the denotation. We should now ask what it means to say that the object in the environment *determines* the thought. Here is one

---

<sup>6</sup> "Other Bodies", p. 98. in *Thought and Object*, ed. A. Woodfield.

<sup>7</sup> As we noted in the beginning of chapter 1, the latter notion is but one of several functions which sense was supposed by Frege to serve.

sense in which it plays that role. Suppose I point to an apple, and express the thought that *that* apple is wholesome. The apple is, as it happens, wholesome, and the *de re* thought is therefore true. Now suppose I point to another apple, that it is phenomenologically indistinguishable, and that I express the thought that *that* apple is wholesome, only despite the apple's outer appearance it is maggot ridden inside: that *de re* thought is therefore false. Hence because each of the thoughts differ in their truth-conditions (they concern distinct objects) the thoughts themselves are to be distinguished. What then if we take the first *de re* thought, and substitute the wholesome apple for another (it too is wholesome)? Given that the truth conditions determine the proposition, and given that in this case the truth conditions are changed, this (third) *de re* thought is also distinct from the first two thoughts. Though this gives us something of the flavour of Burge's singular thought Externalism as suggested in his 1977 paper "Belief De Re", it is only half the story.

For as Burge himself conceded five years later in "Other Bodies", the above illustrates only a trivial kind of Externalism. For few philosophers on either side of the Internalist/Externalist divide would disagree that truth-conditions are external states of affairs and that a difference in the extension of a term is a difference in the truth-conditions of that proposition which the term helps convey. Controversy arises, as we saw in chapter 1, over how truth-conditions themselves are determined - *i.e.* what states of affairs must obtain - for since content determines truth-conditions, how truth-conditions are individuated will depend on how content itself is individuated, which is precisely the issue at hand. Now what makes the above kind of Externalism uninteresting is that even although the propositional attitudes differ in their truth-conditions, they need not - and according to Burge do not - differ significantly in their contents. What do I mean by "significantly"? Firstly, it seems clear that the above *de re* thoughts do differ: after all, they involve different objects. Now as we've already seen, those objects involved in the proposition figure only non-conceptually, whereas what would make for a significant difference is if those three *de re* thoughts of our above scenario differed in their conceptual component; but they don't.

Let us see why. Abstract the context of our three *de re* propositions, and their objects or *res* drop out, leaving us with the *content-clause* (known variously as the propositional function, predicative-fragment, fragment of thought, character, sub-sentential function), which of course is not fully propositional as without truth-conditions it is incomplete or not fully interpreted. Now recall that it is just this aspect of the *de re* proposition which constitutes the conceptual element. The reason why Burge says that it is this the *predicational fragment* or content-clause that is conceptual is because it is this part of the *de re* proposition that captures how the subject himself views the world, a point made clear by the fact that this part of the proposition is expressed in their, what Burge calls, oblique occurrence: to echo what was said above, if "aluminum" occurs obliquely or non-transparently, then "the thirteenth element of the periodic table" cannot enter into the individuation of my belief, even though the terms are coextensive. So were we to individuate mental states in terms of *de re* propositions - in terms of their truth-conditions - then because these aren't, according to Burge, fully conceptualized, there would be no taking into account the way the subject himself views (or thinks about) objects in his environment, leaving us in a poorer position to predict and explain action, and to assess rationality.<sup>8</sup>

So of our three *de re* propositions, all contain the same conceptual element: namely that expressed by "wholesome apple". Hence our little scenario in the paragraph before last only showed contextually different applications of the same content to different entities. For because *de re* propositions are, as noted above, *fundamentally predicational*, the *de re* beliefs can be different at the level of truth-condition, yet the same at the level of content. And insofar as the contents do not differ, nor do the mental states, at least not significantly. In order therefore to get a [significant] difference in mental states, Burge must capture how the subjects themselves view the world; he must individuate mental states in terms of the content-clause of

---

<sup>8</sup> Cf. Chapter 1 above.

propositions, rather than truth-conditions.<sup>9</sup>

**§2.2. The Reducibility of De Dicto to De Re.** It will be instructive if we stay five years ahead, ignoring his "Individualism and the Mental" (1979) for awhile, and focus on what else Burge says in his 1982 paper "Other Bodies". There he says that *de dicto* propositions presuppose their *de re* counterparts. His purpose in saying this - I shall say in a moment what precisely it is supposed to mean - is to undermine in general terms the very idea of "narrow" states which we remarked upon briefly in Chapter 1.

Now we suggested by way of the apple-example that insofar as those mental states in question remained unchanged, Burge's *de re* thought theory of 1977 was not really Externalist at all, that the thesis only becomes interesting where the nature of the mental state itself is claimed to be logically determined by the environment. Indeed it might be thought that the theory of 1977 is in a sense Internalist.<sup>10</sup> After all, we showed that what is common between *de re* and *de dicto* lies in their content-clauses, the conceptual component being precisely what one gets on abstracting the context or environment. In which case one might say that insofar as the content-clauses of our three *de re* thoughts are the same, these content-clauses could be entertained whatever the reality of things happened to be. So it might be supposed that where we do exclude external factors, the kind of proposition entertained is *de dicto* - viz. a content-clause expressed by a closed sentence - and that such propositions are in effect what characterize those "narrow" states mentioned briefly above. For this sort of Internalist is assuming that the kind of generality expressed by *de dicto* propositions - their truth-conditions being specified by the quantifier - is such that not only are these propositions independent from particular contexts, but that, because they are fully conceptual, they are independent quite generally of one's environment. Moreover, one might have thought that, though *de re* propositions are typically

---

<sup>9</sup> Cf. Burge, "Other Bodies".

<sup>10</sup> See G. Segal "The Return of the Individual" *Mind* 1989.

Externalist, we can nevertheless retreat into the "narrow" states presumed to be characterized by *de dicto* propositional attitudes.

But that thought would be wrong, according to Burge. For as said, one of his aims in "Other Bodies" is to establish that, contrary to Internalists, *de dicto* contents are not immune from the external constraints of individuation. And he attempts to establish this by claiming that *de dicto* thoughts presuppose their *de re* counterparts. Now in his 1977 paper "Belief De Re", Burge also aims to establish that *de re* propositions are the more basic than their *de dicto* counterparts. By more basic he means that *de re* propositions are irreducible to *de dicto*: that is to say, where *de dicto* and *de re* proposition have the same content-clause *de re* thoughts are required in order to have the *de dicto* counterpart. And he appears to argue for this by claiming that, unlike *de dicto* thoughts, possession of *de re* attitudes is a necessary condition for using and understanding language, or for any propositional understanding; and for acquiring empirical knowledge. (For example, we need to be shown particular things in the world before we can learn what kind of things they are.)<sup>11</sup>

At this point we might wonder whether Burge was not chasing a red herring. For why should he worry about whether one kind of thought is the more basic than the other? One suggestion is that Burge was worried, at least in the 1977 paper, that if *de re* thoughts are but a species of *de dicto*, in the sense of being fully conceptual, then this would somehow be a concession to the Internalist who, as we've said, might take narrow states to be characterized in terms of *de dicto* contents, where these kinds of propositions are - in a sense later disputed by Burge - independent of any context. (Another suggestion is that he anticipated and wished to avoid Russellian thoughts: I consider this in chapter 6.) For example, Searle holds that all thoughts are fundamentally general - that their contents are fully descriptive or conceptual - and that because of this the environment matters not at all in our having such thoughts as we do. (We will look at Searle in chapter 6.) Nor in "Other Bodies" does Burge drop this idea that *de re* are more basic than *de dicto*. Indeed it is here

---

<sup>11</sup> For a rehearsal of this argument see *Ibid* § ii.

that he insists that *de dicto* thoughts presuppose their *de re* counterparts. True, he is there pursuing the issue about whether *de re* is irreducible in a somewhat different way: to wit, he tries to undermine the idea that narrow states correspond to *de dicto* contents. Yet he had already undermined narrow contents - or at least so the strategy might suggest - three years earlier in the Social Externalism of "Individualism and the Mental". With that strategy at hand, he needn't have bothered about whether *de re* thoughts are more fundamental than *de dicto*, for all our concepts, according to Social Externalism, are in principle externally individuated; so it's not as if the idea of fully conceptualized *de re* thoughts would undermine his Externalism. Seen in the light of "Individualism and the Mental", there's no philosophical mileage to be made, with respect to individuation of content, about which proposition, *de re* or *de dicto*, is the more basic. That then, I suggest, is the philosophical confusion apparent in the development of Burge's thought.

It is true that one purpose behind the strategy of *Other Bodies*, apart from illustrating another way of undermining narrow states, was to propound a Putnamesque essentialism: supposing that *de re* thoughts are essential to certain kinds of learning - so that only what is shared *de re* can be shared *de dicto* - then doppelgangers of different environments cannot share *de dicto* propositions: for each subject incurs a different causal history *via* their different *de re* thoughts. Thus if, as Burge says, *de dicto* propositional attitudes presuppose *de re* attitudes, then it is indeed hard to see how the attitude contents of my twin and I could be the same, for *ex hypothesi*, my doppelganger has never had contact with water, never been related *de re* to it, never been *en rapport* with such a thing; nor has he had contact with anyone else who has, there being no water in the counterfactual situation. And if a causal relation of some sort is a prerequisite to concept acquisition, then nobody in my counterpart's linguistic community even uses a word that means *water*. Our concepts are only superficially similar, albeit strikingly. And to suppose that my twin has a *de dicto* propositional attitude involving the concept "water" is to suppose that many of his ordinary (relational; wide) beliefs are false - *viz.* that *that* is water, that there is water within twenty

miles, that chemists in his country know the structure of water, etc. Hence if *de dicto* presupposes *de re*, these examples will affect narrow propositional attitudes - *viz.* that some glasses somewhere sometime contain some water, that some animals are smaller than all mackerel, and so forth. But of course my twin's propositional attitudes relate to his environment and are, for that reason, by and large successful or true, just as mine relate to my environment are by and large true.

All this might be so, but one still wonders whether the strategy requires of us to say that *de re* propositions are irreducible. Surely it shows, assuming essentialism to be right, only that we need *somehow* to be causally related to the nature of things which our thoughts concern. Again: Burge's Social Externalism should already have settled the issue.

The further moral to be drawn from Other Bodies - one which Burge explicitly draws - is that Putnam's conclusion that psychological states do not determine extension is untenable as it stands. Supposing that *de dicto* propositional attitudes do presuppose *de re* or wide states, then for all Putnam has shown psychological states may well fix extension. After all, psychological states differ just as extensions do. Yet if we maintain, as Burge does, that *de dicto* attitudes precisely do presuppose the existence of entities beyond the subject - and thus that, pace Putnam, "narrow states" so-called are affected by the environment, and are therefore not constant across the relevant possible worlds - then we get the conclusion that psychological states do not determine the extension of natural kind terms. But then if their individuation makes essential reference to the nature of things in subject's physical and social environment, then this notion of a "narrow" state - taken as not having any existential generalization - is dubiously coherent. Moreover, if we follow Burge and individuate beliefs in terms of their "relatively" constant content-clause, and if we claim, as he does, that the content-clauses are themselves individuated by the environment - hence their being only relatively constant - then we must reject Putnam's assumption that our counterfactual man is an exact duplicate in thought, feelings, interior monologue, etc. All said, we must reject Putnam's division between "narrow" and "wide" mental states.

**§2.4. A Non-Linguistic Burgean Thought-Experiment.** In addition to the other forms of Externalism already mentioned, Burge has more recently offered another kind, namely that concerning perceptual content.<sup>12</sup>

There are two visible objective properties O and C. O is a shadow of a given shape and size; C is a crack of similar shape and size. In the actual situation a person (P) perceives a property O in his environment. P is well adapted to the external world and often experiences a [visual] representation of O. In short, O constitutes the content of his experience. Misrepresentation is of course possible; indeed P on occasion perceives C as an O. As such his behavioral dispositions manifest no ability to discriminate C from O.

In the counterfactual situation, person R is P's counterpart: their internal constitutions, *including* behavioral dispositions, are type identical. His environment, however, contains no property of O. Experiences of the intrinsic type T normally produced in the actual environment by O are here normally produced by instances of C. In this counterfactual situation, R has a visual experience of C caused in the normal way by C.

The question then is this: do P and R share the same intentional content? Remember: R's experience of C is intrinsically the same as P's experience when he misperceives a C as an O.

Burge maintains that, though the two experiences in the actual and counterfactual situations are *intrinsically* the same, R does not see the C as an O: so R doesn't have an O-type Intentional state, there being no O in the counterfactual situation. Thus P and R do not share the same Intentional content, nor the same kind of mental states even although their non-Intentionally specified states are the same. Hence, against the Internalist, the local supervenience claim is contravened.

It is time now to assess these forms of Externalism.

---

<sup>12</sup> "Individualism and Psychology", *Philosophical Review*, 1986b; and "Cartesian Error and The Objectivity of Perception", 1986a, in *Subject, Thought and Context*, ed, Pettit and McDowell.

## Chapter 3

---

**§3. A Dual Component Response And the Causal Efficacy of Intentional States.** There is a type of Internalist response, known as the Dual-Component Theory, which accepts by and large the Externalist moral that the contents of propositional attitudes are individuated, in our common sense or every day ascription of them, by facts or states of affairs external to ourselves, and more or less accepts for this reason that a subject's propositional attitudes will differ in their content according as the environment in which that person is embedded differs. But this kind of Internalist response urges at the same time that there is another, more far reaching, moral to be drawn, one which full-blooded Externalists, for their own programmatic reasons (of which more later) deny - namely that we should adopt a more scientifically felicitous method of individuation, a method which, if adopted, would have us revise our ordinary notion of Intentional states, in particular by revising our notion of content.<sup>1</sup>

Now I have been assuming, and will continue to assume, that Intentional states play an explanatory role: that a subject's behaviour is the manifestation of those beliefs and desires which dispose him not merely to act, but to act in a certain way, and that it is in terms of these and other propositional attitudes that action is explained. This assumption then is partly the idea that such states are causally efficacious. And precisely what makes these Intentional states causally efficacious, and what that might mean for individuation, is just the issue we should now explore, for it is this issue more than any other, I think, which makes Externalism seem so puzzling.<sup>2</sup> Indeed all the Internalist responses to be considered in this and the next chapter - and I have room only for Fodor, Segal, and Crane - are motivated by the *prima facie* plausible claim that propositional attitudes, if individuated in that "deep and fundamental" way urged by Externalists, where Intentional contents are logically dependent for their identity on facts and states of affairs in the

---

<sup>1</sup> There are, to be sure, a variety of Dual Component theories: Dennett, and McGinn both share similar sentiments; however, I have room only for Fodor's account.

<sup>2</sup> On the other hand, you might find the epistemological repercussions of Externalism the more puzzling aspect: *i.e.* with respect to knowing one's own mind; but we've not the space for that.

environment - facts or states of affairs whose natures the thinker may or may not be aware - then propositional attitudes cannot figure as explanatory constructs in any theory purporting to be a scientific psychological explanation.

For recall that Externalism violates local supervenience: the view, as Fodor might put it, that there's no difference in one's Intentional states without a difference in one's *physiological or brain* states (*ex hypothesi*, my doppelganger and I are physically indistinguishable, though, because we're embedded in distinct environments, our propositional attitudes are distinct). That might be so much the worse for commonsense psychology. But as far as scientific psychology is concerned, local supervenience is the best idea we've got for showing how mental causation is possible. For in order for there to be a local causal transaction between propositional attitudes and the causal mechanism of the brain, propositional attitudes must be individuated in a way that counts towards mental causation. Otherwise Intentional psychological explanation cannot be preserved.

Crane puts the worry thus:

... the Twin Earth argument entails either that Intentional states are not the causes of behaviour; or that the contents of Intentional states can affect what they cause without any intrinsic causal mechanism. If the first of these is true, then there really is no problem of Intentionality; and if the second is true, then the problem ... is solved by something like an appeal to magic.<sup>3</sup>

The problem of Intentionality just referred to lies in saying how the thinker's intrinsic properties - *i.e.* some causal determinant of action - are correlated with the contents of one's thoughts. For recall that a propositional attitude, taken as an explanatory construct, is simultaneously ascribed both causal and semantic properties, where the former ensures a transaction between mental and physical properties, and where the latter ensures that the action is of the right kind: *viz.* it's the semantically evaluable aspect of the Intentional state such that, for example, if I want a pint of beer then I ask the barman for a

---

<sup>3</sup> See T. Crane, "All The Difference In The World." *Philosophical Quarterly*, 1991. Not that Crane is a Dual Component theorist; of this more later.

beer rather than ask him for his wife. Now one worry is that whereas the causal component of the Intentional state is in the head, *ex hypothesi* the semantic or propositional component with which the state is supposed somehow to be correlated is not. It is hard therefore to see how the two components can conspire to cause action, hard to see how an Intentional state can have a causal role.

The allusion to magic, on the other hand, is the thought that if in fact contents or semantic properties are [somehow] correlated with intrinsic properties - as they must be if the mental states in question are to be both causally efficacious and Intentional - and if those contents are, as the Externalist insists, individuated in a "deep and fundamental" way by the nature of things or facts external to oneself, then the kind of correlation which seems to be implied here is one where, though the contents do indeed conspire to help determine action, they do so without affecting the causal mechanism, and this would suggest that the kind of causation in question involves, in some magical and unacceptable sense - action-at-a-distance. For again, *ex hypothesi* my doppelganger and I are identical in our intrinsic properties. Yet in virtue of being in (minutely) distinguishable environments (though not distinguishable necessarily to the subject) our drilling for "water" (say) counts as different behaviour. Thus our Intentional states (to which behaviour they gave rise) must have had different causal power. Yet there's no trace of that difference on our causal mechanism. So put, the violation of local supervenience seems intolerable. It undermines the very idea of a [scientific] propositional attitude psychology, and makes ascriptions of propositional attitudes look more like a heuristic device useful perhaps for our every day need to make intelligible what people do and say but not something ever to be taken seriously.

In the light of the above objection, the Internalist, whether Dual Component theorist or not, argues that psychological explanation requires that Intentional states be Individualistic, *i.e.* that they be individuated in terms of their causal powers. Fodor<sup>4</sup>, for example, offers the conjecture that science

---

<sup>4</sup> See *Psychosemantics* ch. 2.

itself is Individualistic: that science taxonomizes by distinguishing entities and events by their causal efficacy, and groups things together insofar as their causal powers are the same. So let us ask how it can be relevant to the causal power of Intentional state - to scientific psychology - that water on Twin Earth is XYZ, whereas water on Earth is H<sub>2</sub>O. How can these facts make a *relevant* difference to the way we individuate the causal power of our Intentional states and thus to the individuation of the behaviour to which those states give rise? Individualists reply that they can't. They argue that since *ex hypothesi* our intrinsic properties are the same, and since those properties make a causal difference, our Intentional states must be individuated as the same natural kinds. Hence local supervenience is secured and mental causation is made intelligible, if not explained.

A key question here is thus whether local causation implies local individuation. That it does is one reason for advocating Individualism, and hence for believing in local supervenience. There is, I think, something to be said for this idea; for propositional attitudes do play a role in the explanation of behaviour, and do so in virtue of their causal properties. So how but by respecting the principle of local supervenience between Intentional states and intrinsic properties can those states be causally efficacious? Surely we cannot but run together the notion of individuation and the notion of causation of propositional attitudes.

Burge of course is not unaware of this kind of move, and in his paper "Individualism and Psychology"<sup>5</sup> he raises several points against associating narrow individuation with local causation. He claims, for instance, that the association of individuation with causation involves a conflation. But then he finds not the "slightest conceptual discomfort" in granting on the one hand that causation is local - that physical events in the world affect mental events by affecting local states of the subject's body - and denying on the other hand that the individuation of content is local.<sup>6</sup> Thus Burge by no means says that the

---

<sup>5</sup> *The Philosophical Review*, XCV, No 1 (January 1986) p. 13.

<sup>6</sup> *Ibid.* p. 15-17.

environment can causally affect the subject without affecting one's body. What he actually says is that, although causal variations in the environment can have an undifferentiated affect on the body, the information which affects the body also affects individuation. Looked at in purely non-Intentional terms, my doppelganger and I are indistinguishable; but because we're in different environments, subject to different informational inputs, our attitude states - not least their causal properties and consequences - are different when looked at in Intentional or semantic terms.

This line of reasoning lends itself to the claim that Fodor simply begs the question about what supervenience involves. For what falls out of Burge's principle of [broad] individuation and his claim that causation is local is the idea that mental states supervene on the brain plus environment. In other words, supervenience is a matter of individuation, and individuation is just what is at issue. And that issue, says Burge, can be settled by looking at the actual practice of science. Here he insists that our principle of individuation should derive from the kind of description and explanation which current practice has deemed appropriate to its concerns. And that practice, we are told, is both anti-individualistic and successful.<sup>7</sup> Burge urges that we must, on pain of begging the question about supervenience, avoid criticizing those otherwise successful modes of explanation that happen to rely on external individuation, simply because of one's own metaphysical or a priori presuppositions of what a good entity is or what individuation should be like. Besides, there is, we're told, no superior standpoint from which to judge how (say) psychology ought to be done.

However, the metaphysical point that science must individuate narrowly - that scientific psychology must individuate content in terms of what's causally relevant - is surely sound. The association of local supervenience with causation - whatever science gets up to - are core assumptions of causation and the casual status of mental states. And what is causally relevant dictates the principle that content supervenes locally on intrinsic properties.

---

<sup>7</sup> Cf. Burge, *op. cit.*, p. 16-17.

At any rate, Fodor isn't *a priori* about the *sufficient* conditions which count as good explanation; he's postulating only a necessary condition, viz. local supervenience. Fodor's reply then would be that we closely associate causation and individuation on pain of the consequence, suggested in the above quote, of not doing so; so there's no conflation; nor any question begged.

As for scientific praxis being non-Individualistic, it is more likely that the relational properties in question - relations which figure in individuation - are ones affecting the causal powers of the state or entity, and this, Fodor says, is just what Individualism captures:

"... being a planet is a relational property par excellence, but it's one that individualism permits ... For whether you're a planet affects your trajectory, and your trajectory determines what you can bump into; so whether you're a planet affects your causal powers ..." <sup>8</sup>

So a planet's environment contributes to - and of course does not constitute - its causal power, like its velocity, and the way it moves. Indeed the planet's very nature depends on facts about the environment, since the environment disposes the object to act in a definitive of being a planet.

On the other hand Stalnaker suggests in his paper "On What's in the Head" that individuation by way of what affects causal powers is too weak for what the Individualist wants, and that Fodor therefore conflates this [apparent] weaker claim with the stronger one of individuating by causal power.<sup>9</sup> Individuation by what *affects* causal power is too weak, says Stalnaker, because this is not in any way incompatible with how Externalist individuates mental states. An external fact like, *there being water in the environment*, he says, may affect the causal power of my internal states, viz. by putting me in an internal state that disposes me to behave in a certain way. The compatibility is clear, says Stalnaker, from observing that though Twin Earth arguments show that this same internal state can be had by a subject with a different causal history, much the same can be said of the planet example: we

---

<sup>8</sup> Op. cit., p. 43.

<sup>9</sup> In *The Nature of Mind*, ed. D. Rosenthal, p. 586. (Oxford University Press) 1991.

are asked to imagine a Twin Earth that is not a planet and then to imagine it in an environmental configuration exactly the same as our planet is in. That Twin Earth then becomes a planet supposedly shows that Fodor is closer to Externalism than he wants.

But Stalnaker's point, it seems to me, is weak. As should already be apparent, it certainly doesn't show that Fodor conflates individuation by causal power with individuation *by what affects* that power. And it certainly doesn't mean that this "weaker" claim is compatible with Burge's Externalism. For Fodor argues in *Psychosemantics* that the causal powers must be individuated in terms of the truth of their counterfactuals, that the identity of causal power is identity of causal consequence across nomological possible worlds. Thus my mental state which concerns H<sub>2</sub>O has the same causal consequences as my doppelganger's mental state which concerns XYZ because, though in your context the mental state causes the behaviour of drinking XYZ rather than H<sub>2</sub>O - whereas my mental state causes the behaviour of drinking H<sub>2</sub>O - the following counterfactuals are true: namely that if my doppelganger's utterance or thought had occurred on Earth it would have had the effect that my utterance or thought had had: and *vice versa, mutatis mutandis*. Similarly, *mutatis mutandis*, for Fodor's Planet example: for his criterion of identity for causal power applies as much to what affects causal power as to what constitutes that power. So individuation by what affects causal power is not a weaker claim, much less a claim that is compatible with Externalism.

However, what Stalnaker should have said is what Fodor himself admits: that the above way of putting the point, though sound in spirit, is not quite right. Though Fodor doesn't say as much, the most telling passage in which the error occurs is, I take it, the following, particularly in the second and third sentences:

"This line of argument ... assumes that I can say "Bring water!" on Twin-Earth ... But although I've heard it suggested that mental states construed nonindividualistically are easily bruised and don't "travel", the contrary assumption would in fact seem to be secure ... That if, standing on Twin-Earth, I say "That's water" about a puddle of XYZ, then what I say is false. Which it wouldn't be if I were speaking English<sup>2</sup> ... So ... we have ...

no reason to suppose that the causal powers of my Twin's are different ..."<sup>10</sup>

The paragraph is a give away; it invites an Externalist reply that since if on Twin-Earth I utter "That's water" and it is false, then equally if I say "Bring me water", where my doppelganger on Earth says "Bring me water", then, given that our Intentional states travel (they're not easily bruised), neither of us get what we want, for you get H<sub>2</sub>O and I get XYY.

Fodor has since reformulated his intuitions, however.<sup>11</sup> He agrees that myself and my doppelganger differ in our causal properties, but rejects, as he must, that this difference is one of causal power, at least not in the sense of its being responsible for a certain difference between their effects, *i.e.*, where one effect has a property which the other effect lacks. This is not to deny that having water thoughts and twater thoughts have causal powers, only that those powers are distinct in the relevant way. For the causal power of our thoughts are simply instantiated in subjects whose different causal histories accounts for any property differences they may have. And that these differences in causal history don't count as differences in causal power is because it's *conceptually necessary* that if I'm connected to water then I have water-thoughts, for to have a water thought just is to have a thought connected to water. So in order for the connection between having a water-thought and the consequent behaviour to count as a difference in causal power, that causal power must enter into non-conceptual (*i.e.* contingent) relation. So for example, if E is a property that Keplerian orbits have then, if x is a planet with a Keplerian orbit, x's orbit has E. And it is contingent that if x is a planet then x has the properties of being a Keplerian orbit, for it's contingent that planets have Keplerian orbits.<sup>12</sup>

---

<sup>10</sup> Op cit. pp. 35-36.

<sup>11</sup> See his "A modal Argument For Narrow Content", in the *Journal of Philosophy* 1991.

<sup>12</sup> You might claim, however, that B is the property that my water thoughts are thoughts about y's favourite drink, and that this is clearly contingent. But that still won't count as a relevant difference to psychological taxonomy: though B is a property which water thoughts have contingently, it is conceptually necessary that if I'm appropriately connected to water then my thoughts are water thoughts; and equally it is a conceptually necessary that if something

**§3.1. The Very Idea of Narrow Content.** What, now, of his proposal to revise content: that revision which for methodological reasons Intentional psychology must undergo?

The point made in Chapter 1 that content is a sufficient condition for determining its extension was important not least because truth-conditions are themselves individuating of thoughts or propositions, and because extensions are partly what truth-conditions are defined over. But on Fodor's story it seems that though on the one hand the thoughts of myself and doppelgänger must, given the local supervenience constraint, be the same, on the other hand given his Semantic Externalism, these thoughts have different extension, so that my thought can be true where his is false. But how can our thoughts be both the same and different? What is the thought such that when I have it its truth-conditions are that H<sub>2</sub>O is wet, whereas when my doppelgänger has it its truth-conditions are that XYZ is wet? It looks indeed as if the sameness of mental content is no longer a sufficient condition for determining the sameness of its extension. Yet unless we maintain that the identity of content *is* determined by the identity of their extensions - unless we say: *no difference in extension without a difference in content* - psychology cannot say what counts as evidence for the same propositional attitudes. Supervenience of mental states on physiology may have facilitated the intelligibility of how propositional attitudes can play a causal role, but only at the price it seems of vitiating the principle that contents supervene on their extensions - and with it, perhaps, the intelligibility of propositional attitude psychology.

To Fodor, however, this is not so. In order for extension to constrain the identity of Intentional states, we must, he urges, relativize the connection between content and extension to context, so that content still determines extension, but only relative to context. To see what relativization means, let us look briefly at Kaplan's (1977) account of the semantics of indexicals, the inspiration behind Fodor's revision of content.

---

is a property of water thoughts - like (B) being a thought about y's favourite drink - then it is a property of water thoughts if my thoughts are water thoughts. Cf. Ibid.

The relativization with respect to indexicals is easy to see. Take the sentence "I'm hungry" as uttered by myself and my doppelganger. The connection between content and extension, that connection which gets us to a fully propositional content, is not constant but relative to the context (here the person whose utterance it is), this because context contributes to determining the truth-conditions. Thus the difference between your use of the sentence "I'm hungry" and my use of the same sentence is that its *character* and use determines the truth-conditions which, being a function of context, yield distinct propositional contents.

Despite the difference between the proposition expressed by my own and my doppelganger's utterance of "I'm hungry" there is clearly something that is the same. That sameness, we're told, lies with what remains when we abstract the context; and it is this "sameness" between us which Fodor takes as his new notion of content. What remains, at least according to Kaplan's theory, is something like an the open sentence or sentential function whose meaning, fixed by convention, is constant across context (its *character*). And the same it must be if the *character* of the open sentence is to determine the context, and with it a fully propositional content: something that is true or false according as the person who utters the sentence has the property of hunger.

Now Fodor takes this idea of an open sentence with a character as its constant component, and generalizes its application beyond the semantics of indexicals and demonstratives. It is for him a new kind of content, a narrow content, a function which takes context onto truth-conditions: "narrow" because it is intrinsic (in the head) and - like the *characters* of indexicals - is [apparently] not context-relative. And what makes it "content", according to Fodor, is that it determines, relative to context (and given the meaning or *character* of the sentence), the semantic or Intentional properties of the thought or proposition. So what myself and my doppelganger have psychologically in common, seeing that we're physiologically identical, is something like an open sentence  $f(x)$  such that when I think "water is wet", this internal state takes or determines the [Earthly] context in which it is uttered as argument and yields a broad content, a fully truth-conditional thought as value. And similarly for

my doppelganger. The sentence "water is wet" uttered by him has the same narrow content yet determines his context (*i.e.* Twin Earth) - takes it as argument - and yields or determines a broad content or fully truth-conditional proposition as value. In a word, narrow content is a function from context to truth conditions.

We should note that this conception of narrow content conforms to his picture of a language of thought.<sup>13</sup> This is his [empirical] hypothesis that the brain computes in a language - a language which mirrors thought - whose symbols or "representations" are distinguishable by their shape (like letters *vis-a-vis* words), and whose individuation is by their functional or causal role. In having the same proximate stimuli, and in being physiologically identical, my twin and I token the same syntactic objects "that water is wet" whose *character*, though also the same, will determine, given the context and truth conditions, different broad contents or propositions.

So where we can say that a propositional attitude is a relation between a person and a proposition, we can see that, given Fodor's picture of a language of thought, propositional attitudes collapse into the relation between a person and a mental sentence, and the relation between a mental sentence and the proposition expressed. Moreover, though our distinct contexts help to yield distinct broad contents, we can still say, in virtue of sharing the same narrow contents (that being the content), that we share the same Intentional states. The semblance of the relation between content and believer is thus preserved, as is the structure of Intentional explanation; or so we might suppose.

Whether or not there is a language of thought is of course not the issue here. Indeed as a general hypothesis it is not incompatible with anti-Individualism (so long as semantic content doesn't supervene on syntax). I mention it partly to show Fodor's motivation for advocating narrow content - namely as being the causal component of content - and to show further why he believes that wide-content is irrelevant to psychological explanation. But

---

<sup>13</sup> Cf. Fodor's "Methodological Solipsism Considered as a Research Strategy for Cognitive science", in *RePresentations*.

I also mention it because it is a wonder that we should ever take pieces of language of thought as having content independently of context and truth conditions. Why believe that the function is [narrow] content? And how in any case are we supposed to identify that function?

Let us take the last point first. For whenever you assert what the thought is, it becomes "anchored" to the context in which it is asserted, so that by acquiring a semantic value the content fails to be narrow, making it a wonder that we can describe what this content is that my doppelganger and I are said to share. Looked at this way, narrow content appears mysterious: mysterious because ineffable. As Fodor himself says:

"Narrow content is radically inexpressible, because it's only content potentially; it's what gets to be content when - and only when - it gets to be anchored. We can't ... say what Twin Thoughts have in common. This is because what can be said is *ipso facto* semantically evaluable; and what Twin-thoughts have in common is *ipso facto* not".<sup>14</sup>

The remark that "narrow content" is only content potentially inspires our first question: *i.e.* why believe that what is "narrow" is content. At any rate, is it not odd to identify things in terms of what potentially they are: we're all potentially insane, but that doesn't make us mad. It is one thing to say that our states have syntax; but why believe they have content at this level, not least when [apparently] it can't be identified? We'll return to this.

For the question currently under consideration is the second one, how we might identify that function which effects the mapping from context onto truth-conditions. Now Fodor suggests that we can "sneak up" on it: we can mention rather than use a sentence and in this way pick-out the narrow content of that mental state which myself and my doppelganger share: *i.e.* by seeing what broad content my doppelganger would have were he, as Fodor puts it, "plugged into my context" or my world. To do so, however, we must first abstract the context or relational properties since there is, says Fodor,<sup>15</sup> a

---

<sup>14</sup> *Psychosemantics*, p. 50.

<sup>15</sup> *Op cit.*, p. 48.

relation between Twin-Earth and Twin-me (condition C) in virtue of which my doppelgänger's "water"-thoughts are about XYZ, but mine aren't. Still, were I to share condition C, then given that we're doppelgängers, our narrow contents would *effect the same mapping of contexts onto truth-conditions*. The assumption is that "context C" contains all the "relevant information" external to the believer, such that context together with the internal sentence of the believer will determine the truth conditions.

Stalnaker suggests that this is banal: for we knew already that "truth conditions are determined by the conditions that are relevant to determining them";<sup>16</sup> which is all that condition C tells us. That is, we already knew that given our internal state - e.g. the internal sentence "water is wet" - its character together with context determine the truth-conditions. But what we want to know, and what Fodor doesn't say, is what that *function* narrow content is supposed to be.

To illustrate, Stalnaker offers the following parody of Fodor's way of narrowing content, the idea being that if Fodor's criterion worked it could be used to narrow the content of any relational property. Suppose I have the property of being three miles from x (condition C) but you don't; instead you, my doppelgänger, have the property of being three miles from y (conditions C\*). Now Stalnaker says:

"Whatever these conditions are, we do know this: short of a miracle, it must be true that anyone in the location that both I and my counterpart are in in our respective worlds would be three miles from x if conditions C obtained, and three miles from y if condition C\* obtained".<sup>17</sup>

Stalnaker is here complaining that this doesn't help us identify a specific function that takes condition C, the context, into the property of being three miles from x - that specific function being the contribution my location makes to my having the relational property. In other words, Fodor's procedure is too obscure. Stalnaker again:

---

<sup>16</sup> Op. cit. p. 580.

<sup>17</sup> Op. cit. p. 580.

"There are many such functions, and no reason to identify any of them with the contribution that my intrinsic location makes to the specific relational property ... Every location is such that for some external condition, if those conditions obtain, then anything in that location is three miles from a burning barn".<sup>18</sup>

You might ask whether the parody is not misplaced. For why should we expect Fodor's identity criterion for narrow content to yield the goods for every relational property? Is Stalnaker's trivialization the result not so much of a failed criterion but of applying that criterion to an irrelevant example? After all, who would expect Fodor's procedure to identify an explanatorily useful function of one's location?

However, the crux of Stalnaker's point is, to be sure, a serious one, casting doubt on whether a useful function can be identified. Imagine scientists trying to identify the internal sentence or narrow belief that, e.g., *salt is soluble in water*, independently of its wide-content, i.e., *salt is soluble in water*, as a sentence in the brain. Of course they might succeed. But what if the content is determined holistically? Then that belief-state would consist of complex clusters of dispositions to behave in various ways under various conditions. True, the "total belief" state might nevertheless be described by the particular brain-sentence, "salt is soluble in water", so that the truth that I believe the proposition *salt is soluble in water* would be shown by describing how, given the external context, I would be disposed to behave under various conditions. But equally the complexity of that "total belief state" might make it impossible to match the *particular* belief-content with the *particular* dispositions in the cluster. For other dispositions of the same complex cluster might make it true that I believe a number of non-equivalent propositions. To use our above example, the answer to the question, what makes it true that, given conditions C, I'm three miles from x, will be answered by describing the factored-out component of my relational property, i.e. my location. And Stalnaker's suggestion is that this intrinsic component won't be explanatorily useful, for the same function will also make it true that I've lots of non-

---

<sup>18</sup> Op cit. p. 580.

equivalent relational properties, say, being an inch from an ant, two days from the moon, being closer to the ant than home, and so forth.

The trouble then is that Fodor doesn't tell us what to factor out. And unless we can be clear about what to factor out, we'll have no criterion of identity for narrow content. And unless a function can be identified as being the same or different, such a theoretical construct will be useless in any (so-called) Intentional explanation. The problem is still more vivid with respect to our comparison of ordinary indexical expressions. The sentential functions "I", "here", "now" are defined over their contexts types which, as it happens, are well-demarcated: respectively, the person, the place, and the time. For that we can distinguish the type of indexical involved - the sentential function - is itself a function of the type of context that the indexical takes as argument. But well-demarcated is precisely what the context for narrow content is not. In short, we don't know what function narrow content is supposed to be.

What, now, of Fodor's remark that narrow content is content potentially? There appears no reason to suppose that there is an analogue for *character* - one that is constant across contexts - which determines the mapping of contexts onto truth-conditions. It is one thing to say that myself and my doppelganger share the same internal state in the sense of having the same causal power, and that such states for this reason can be taxonomized for purposes of scientific psychology as the same natural kind. It is another thing, however, given that we have no clear notion of narrow content, to say that this common state can be characterized in anything like Intentional terms. After all, (and to paraphrase Fodor) content isn't content unless constrained by a criterion of identity.<sup>19</sup>

---

<sup>19</sup> Cf. op. cit. p. 46-47.

## Chapter 4

---

**§4. Defusing the Twin Earth Thought Experiments.** So where does all this leave us? It was suggested in the last chapter that the association of individuation with causation might be a good idea: that sound methodology requires of us to type-individuate the Intentional states of myself and my doppelganger. But as just seen, the Dual Component theory leaves us in a bind: accepting semantic Externalism forces us to construct a theoretical notion of content in order to accord with (i) local supervenience - *i.e.*, the prerequisites of local causation - and (ii) the prerequisites of Intentional explanation - *i.e.*, that the model of explanation involve a notion of content. You might conclude therefore that since the idea of "narrowing" content is found wanting, the moral of the Thought-Experiments is that scientific propositional attitude psychology is untenable.

That conclusion, however, would be premature. For one might accept the principle of individuation of content by way of causal power yet reject Dual Component theories. But that means the rejection of Externalism *tout court*. Such is the kind of strategy which Crane, from whom we quoted above, deploys. This chapter will try to show how we might collapse the Twin-Earth arguments in order to put Intentional psychology, given our stricture on sound methodology, back in the running. So let us see how we might dissolve the Thought-Experiments, taking Putnam first.

Why in any case believe that the chemical structure of water affects the truth-conditions of the sentence (say) "water is wet" in the way Putnam says? For you could say that a discovery of XYZ would simply confirm what we already know from our discovery of isotopes: that not all water has the same microstructure; that just as  $D_2O$  is as much what we call "water" as  $H_2O$ , so too is XYZ. We could say that since water can be realized by a multiple of microstructures, the concept would have a disjunctive content, *i.e.* water is  $H_2O$ , XYZ,  $D_2O$ , PTX, KYZ, etc, where the latter two are newly discovered microstructures which realize or instantiate water.

As for his other examples, it is all very well for Putnam to say that when the names and (relative scarcity) of aluminium and molybdenum are

switched on Earth and Twin Earth the meaning of "aluminium" spoken by my doppelganger and I is underdetermined by our narrow psychological states. For seeing that we're unable to distinguish between aluminium and molybdenum, our narrow states individuate aluminium *and* molybdenum as the extension of mine and my doppelganger's use of "aluminium". But it's quite another thing to say that my use of the word "aluminium" expresses the concept molybdenum, whereas my doppelganger expresses the concept of aluminium. After all it's not what our linguistic behaviour would suggest: namely that neither of us fully understands the meanings of our words. In which case why not say, as Crane<sup>1</sup> does, that we share the same concept (Crane calls it "molyminium") whose extension includes both metals? The difference between the concept of *molyminium* and the concepts *aluminium* and *molybdenum* is that the former distinguishes less finely than do the others: though it distinguishes the two metals from other metals, it doesn't distinguish between the two metals themselves.

It would be mistaken, I think, to conclude that meaning now no longer determines extension. Such would be so only where the extension of mine and my doppelganger's use of the word "aluminium" are different. But arguably they're not: our common concept *molyminium* determines the extension that includes both metals; and moreover, it is in virtue of this that we manage to refer in our respective environments. Of course were we to have complete understanding then, given communal standards, the concept *aluminium* would determine the extension aluminium. But it's not complete. And that my doppelganger and I do not understand the [public] word "aluminium" completely is hardly a reason to deny that we altogether lack understanding; for it is quite plausible to distinguish, as we have, between the public meaning of a word and the concept that one intends to express.

**§4.1. Criticisms of Burge's Social Externalism.** What, now, of the alleged error in Alf's belief? (Recall that according to Burge our Intentional states

---

<sup>1</sup> Tim Crane, "All The Difference In The World", *The Philosophical Quarterly*, 1990.

differ according as the use of our words differ.) There are a number of possibilities. Either Alf is mistaken in his belief and therefore in his concept; or he is mistaken in his use of language; or both. Or perhaps he is mistaken in none of these things.

Crane suggests the second option. When Alf expresses his belief with the sentence "I've got arthritis in my thigh", that belief is expressed *via* a second order [general] belief (or habit) about whether the words are the right ones to express the concept. So according to Crane, Alf's error lies not in the belief that (i) he's got arthritis in his thigh but in (ii) that "I've got arthritis in my thigh" is the right sentence to express this belief (where in (ii) "this belief" refers to whatever belief Alf has). And clearly if Alf believes that "I've got arthritis in my thigh" is the right sentence to express his belief, then by communal standards his second-order belief (*i.e.* about his language) is false. But this needn't mean that his error lies also in (i). That depends on whether Alf was expressing with the word "arthritis" the public concept arthritis. Yet Alf may have deployed a private concept, say, *tharthritis*, whose extension includes both arthritis and whatever ailment he's got in his thigh. So the belief involved in (i) would be, I've got tharthritis, and would be true; whereas the belief involved in (ii) would be false.

How do we decide whether Alf's error is linguistic or conceptual, or both? As Fodor points out<sup>2</sup>, being unable to distinguish in experimental situations between the non-verbal behaviour of Alf and his doppelganger, a psychologist would not hesitate to ascribe a common reference. This is because the subjects have the same dispositions; dispositions which furnish them with no ability to discriminate between the extensions in question, where the latter skill is surely a prerequisite for the possession of a different concept.

And about this [metalinguistic] strategy itself, there is, it seems to me, nothing strange about it. Burge, though, is unconvinced; he claims we can make good sense of someone who says, for example,

( $\beta$ ) For years I believed that a fortnight was ten days, not fourteen, though of course I never believed that fourteen days

---

<sup>2</sup> *Psychosemantics*, p. 40.

were ten days.<sup>3</sup>

without having to attribute beliefs about the meaning of the words one uses. He also holds that attributing metalinguistic beliefs isn't in any case our ordinary way of interpreting what people say.

It is true that we can make sense of the thinker in question, and that we would not attribute contradictory beliefs to him. But the question is how would we make sense of him? One suggestion then is that we attribute to him an erroneous belief that "fortnight" means a period of ten days. However, Crane goes so far as to deny the possibility that we could express our thoughts in words and not have beliefs about them. And in any case, it is hard to see, he says, what else "fortnight" could have meant to him.<sup>4</sup> In support of this [strong] metalinguistic thesis he has us envisage what the speaker would say if asked whether he believes that "fortnight" means a period of ten days. And of course the speaker would answer "yes", "no" or "I'm unsure", thus articulating a belief about the meaning of a word. But this is unclear: equally it can suggest that the speaker has a metalinguistic belief only on being asked the above question; the question forces him, as it were, to have a belief which hitherto he might not have had.

This brings us to the last option. The Internalist can dispense with the metalinguistic strategy, by arguing that the speaker is mistaken neither about his language nor about his concepts: that is, the Internalist needn't attribute *any* false beliefs to Alf, neither about his choice of words, nor about the concept used. This is because the speaker's self manufactured concept can define the use of his word, and with it its meaning.

How good then is this strategy of concept attribution? It will be pointed out that though the attribution to the subject of a private rather than [false] public concept succeeds in disarming Burge, the strategy comes at a price not worth paying: in short, that it eliminates the possibility of conceptual error. Whether or not Alf has a false belief about arthritis, it is true that he

---

<sup>3</sup> T. Burge, "Belief and Synonymy", *Journal of Philosophy*, p. 126. 1978.

<sup>4</sup> Op. cit. p. 20.

uses a concept in a systematic way: it being precisely because the subject has this ability to classify things regularly over time that we attribute a concept to him at all. Now it is because the concept is here deployed to classify a thing regularly over time that I feel compelled to say that the subject is not in error about the concept he has. And this means we should rule out what might to you seem equally compelling: the possibility of conceptual error *per se*. For whenever the subject's use of a concept falls short of the public standard, the deviant subject is seen simply to be reclassifying the world. And provided that reclassification is systematic, we can say that his use of the concept, and thus the very content of his belief, is true.

So I suggest the Internalist should bite the bullet: the idea of conceptual error, of the kind outlined, is a misnomer. There is nothing absurd about this claim; nor does it in the least impugn the plausibility of the private concept strategy. Indeed it is part of its strength. That anyone should think otherwise is a consequence of their "misconception" of what kind of error might be in question. One bad reason for supposing that my suggestion is absurd is this: that immunity from conceptual error means that if a subject can never be wrong about the use of a public concept, neither can he ever be right, since getting things right depends on the possibility of getting things wrong. In which case, a subject can't have public concepts. But in fact this immunity from error thesis does nothing to preclude our use of public concepts. For what we should say is that either you grasp the publicly defined concept or you do not; not grasping it, however, hardly constitutes getting it wrong, still less does it preclude the possibility of apprehending it in the future.

You might ask whether the possibility of error *tout court* is here being ruled out, or whether the thesis applies only to public concepts. Might error with respect to the deployment of one's own concepts be possible? You might, for example, have a series of temporary memory lapses, where the series consists in a repetition of forgetting and remembering, such that no longer do you apply the concept in the same systematic way. Instead you simply reclassify things, not knowing that hitherto those things in the world were grouped together differently. Then on regaining your memory, realizing

that things have been differently individuated, you might judge that you were wrong. Of course on the other hand you could be your own judge: you might simply settle on the new classification, allowing yourself to have been right on both occasions.

Lastly, and perhaps most importantly, it should be noted that I have ruled out only the possibility of systematic error, and not, so to speak, one-off error. It is undeniable that the latter kind does occur, and that it happens quite a lot. We often do things, or say things, or think things, only then to correct ourselves, or allow others to do so. But that we are liable to be corrected about our use of public concepts is only because it is a one off error or slip. Matters are quite otherwise, I have urged, where the deviation becomes systematic.

In sum, despite my doubts about his metalinguistic argument, the soundness of Crane's private concept strategy remains; and though, in particular, it would be fatuous to deny the possibility of one-off error, attributing systematic error might now strike you as a question-begging thing to do.

With this alternative notion of narrow content in mind, let us finish this chapter by show how Burge's final argument for Anti-Individualism (his non-linguistic strategy touched on in chapter 2) can be unsaddled.

#### **§4.2. Burge's Non-Linguistic Thought Experiment: Perceptual Content.**

(i) Is then your *seeming* to perceive x individuated, as Burge says, by a contextual (a causal or perhaps evolutionary) factor with x, where "seeming" is a conscious state of one's phenomenology (*the something it is like to be in state*)? (ii) Or is that *seeming* type-individuated, as the Individualist insists, by proximate stimuli and behavioral dispositions?

The stakes, recall, are these:

If (ii) is true, then perceptual content supervenes on one's internal constitution: *i.e.* no difference in one's phenomenology without a difference in one's internal state. But if (i) is true then (ii) is false, and anti-Individualism, *vis-a-vis* perceptual content, is true.

**§4.3. Burge's Thought-Experiment.** Take two visible objective properties O and C: O is a shadow of a given shape and size, C a crack of similar shape and size. In the actual situation a person (P) perceives the property O (shadow) in his environment. Because P is well adapted to his world and often experiences a [visual] representation of O (shadow), P's experience is ascribed the content O. Moreover, P sometimes misperceives C (crack) as a O (shadow), a symptom of which misrepresentation is that his behavioral dispositions don't discriminate C (crack) from O (shadow).

Now person R is P's counterpart: their internal constitutions, *including* behavioral dispositions, are type identical. His environment, however, contains no property O (shadow). Experiences of the intrinsic type T normally produced in the actual environment by O (shadow) are here normally produced by instances of C (crack). Also, R has visual experiences of C (crack) caused in the normal way by C (crack).

So do persons P and R share the same Intentional content? (Remember: R's experience of C (crack) is intrinsically the same as P's experience when he misperceives a C (crack) as a O (shadow).)

Burge thinks not. Even though the actual and counterfactual experiences are *intrinsically* the same, R doesn't see the C (crack) as a O (shadow). For there simply is no shadow (O) in the counterfactual situation. Moreover, Burge claims that all this accords with actual practice in psychology.

**§4.4. An Internalist Response.** So what content does R have? In the first instance Burge says that R has intentional state C,<sup>5</sup> then waives, claiming that dialectically it matters only that R and P plausibly have different contents.<sup>6</sup> In other words, Burge doesn't say.

About R's content, an Internalist, on the other hand, is clear. For he

---

<sup>5</sup> T. Burge, "Cartesian Error and the Objectivity of Perception" In Pettit and McDowell (eds), *Subject, Thought, and Context*, Oxford: Oxford University Press, 1986.

<sup>6</sup> "Authoritative Self-Knowledge and Perceptual Individualism". In R.H. Grimm and D.D. Merrill (eds.), *Contents of Thought*, Tuscon, AZ.: University of Arizona Press, pp. 86-98.

urges that we be guided by the subject's behavioral dispositions. And though the distal causes differ - one is a crack, the other a shadow - the experiences themselves represent the world no differently to the agents, and for this reason furnish the subjects with no discriminatory ability. So P and R must share the same perceptual contents.

For suppose that whether something is a shadow or a crack is important for the adaptive success of a creature. Shadows might be their only source of shelter from the mid-day sun, whereas the cracks might be large enough for the creature to risk injury should it fall into one. And if a creature repeatedly does just that, we surely wouldn't say that it perceptually represents the cracks as cracks. We would cite shadows.<sup>7</sup>

Of course, the type of experience produced by the environment might not involve adaptive behaviour. Gabriel Segal<sup>8</sup> has us imagine that in the actual situation, in contrast with normal circumstances N, (wonky) circumstances W are abnormal and non-ideal, and that a C (crack) is misperceived as an O (shadow). Furthermore, P sees O (shadow) in N as O (shadow). In the counterfactual situation, however, wonky circumstances W are normal, and R perceives C (crack) as C (crack). Whereas if N obtains (*i.e.* what in the actual situation we count as N), and instances of O (shadow) occur, R would misperceive O as C (crack).

Now suppose, says Segal, that persons P and R are involved in two experiments in the same laboratory:

Experiment 1: C (crack) is presented to P and R in W. Hence: P misperceives C as O (shadow); R perceives C as C.

Experiment 2: O (shadow) is presented to P and R in N. Hence: P perceives O as O; R misperceives O as C (crack).

Now because neither subject is able to discriminate between O in N and C in W, no behavioral differences between P and R will be discerned. Hence,

---

<sup>7</sup> Cf. R. J. Mathews (1988): "Comments [on Burge 1988a]" in *Contents of Thought*, R. H. Grimm and D.D. Merrill, eds., Tuscon, AZ: University of Arizona Press, pp. 77-86.

<sup>8</sup> Cf. G. Segal. "Seeing What Is Not There". *Philosophical Review*, 98, pp. 189-214.

says Segal, scientific psychology must attribute the same content to P and R's experience: in particular, a concept whose extension includes shadows and cracks, a "crackdow".

**§4.5. An Externalist Response.** But what of P's ability to discriminate C from O in N? The content of P's experience caused by C in N is distinguishable, on behavioral grounds, from the content of his experience caused by an O in N. Yet this fact is blurred by the ascription of the concept *crackdow*, and with it one's intuitions about objectivity.

Martin Davies offers the Internalist two options for explaining the discriminatory ability such that if we reject the first, as he believes we must, we should accept the second which, as he puts it, is something of a poisoned pawn.

*The first option.* P in N perceives C (crack) as a C (crack), and perceives O (shadow) as a O-or-C. But if the Intentional description C of P's experience is a subclass of O, where the content crack is the content crackdow, we lose the sense in which P is able in N to distinguish C from O. For *ex hypothesi* a crack is a crackdow. (Imagine distinguishing crackdows that *are not* cracks from crackdows *that are* crackdows. It would involve seeing shadows as crackdows that are not cracks: as shadows, in other words. Hence we defeat the object; we're left with different contents across duplicates.)

*The second option.* Here we admit that the experiences differ intrinsically. And though the difference is explained by disjunctive contents,<sup>9</sup> it involves reference to environmental circumstance: *e.g.*, for representations caused by C in N we have a *crack-in-N-or-shadow-in-W*; for representations caused by O in N we have a *shadow-in-N-or-crack-in-W*. Hence the discriminatory ability is made sense of, for cracks in N are represented differently from the shadows in N, even although we attribute the same content

---

<sup>9</sup> Note that "disjunctive" is Davies' way of characterizing crackdow. Segal (1991), on the other hand, sees a crackdow as a concept which includes the extension of a crack and a shadow, and not a concept which includes a crack or a shadow; the importance of this distinction will become clear below.

to P and R.

The "poison" of this option comes to light, however, when we spin a further thought-experiment.

Suppose experiences of intrinsic type T, normally produced by O (shadow) in the actual situation, are normally produced by C (crack) in the counterfactual situation. Moreover, because of peculiar optic laws operating in the external environment of the counterfactual situation, experiences of intrinsic type T, produced by C in N, are not produced by C in W, and nor are they produced by O. With what content, then, do we credit R's experience?

Clearly we cannot ascribe to R what we ascribed to P's actual experience of the same intrinsic type T: namely the disjunctive content *O in circumstances N or C in circumstances W*. For R's experiences are never produced by O; they're normally produced by C in N: *i.e.*, not by C in W (imagine that the experience is explained by an optic law peculiar to that environment). Thus this ascription of the same disjunctive content to R and P would blur the normal circumstances in which experiences are generally veridical. Nor can we attribute to R's experience the content C in circumstances N. For perceptual content would not then be preserved across duplicates.

The only alternative - and here the "poison" takes its toll - is to expand the disjunction. R's counterfactual experiences are, we saw, of the same intrinsic type as P's actual experiences, namely C. Now that will warrant at least three disjuncts. And each disjunct will have to mention the objective properties (e.g., O and C), the environmental circumstances (e.g., whether N or W), and other background conditions (e.g., the optic laws). Nor is that the end of it. For in response to additional disjuncts, the Externalist will exploit the points about objectivity and the possibility of error by spinning still further counterfactual arguments.

The consequences for Intentional explanation are stark. Recall that the intrinsic or phenomenal character of experience has been equated with intentional content, that the disjuncts refer to external circumstances, so that, at least in normal circumstances, any difference in intrinsic experience

corresponds to a difference in the distal cause. So when specifying the difference in intrinsic experience, one of the disjuncts will be different. Now when we speak of phenomenology we speak of how things seem; we speak of the *possible* states of affairs as the cause of the experience, and thus as explanatory of behaviour. Thus we cannot use such a straightforward content as, *e.g.*, that it seems to the subject that there is a cube four feet in front of her. For we are straddled with indefinitely long disjuncts "seemings". And clearly these indefinitely long disjuncts cannot figure in the Intentional explanation.

**§4.6. The Status of the Thought-Experiment.** Thus the "poison" of the pawn takes its toll on spinning further Externalist Thought-Experiments. But to the Internalist such spinning of counterfactuals can't get off the ground for they rest on the implausible premise that the counterfactual subject is successfully adapted to his environment. For consider: In the actual situation, P in N sees a cube four feet in front, which causes representation of syntactic type S, giving S the content, *there is a cube*. In the counterfactual situation, where N is rare, R in W sees a sphere four feet in front, which causes representation of syntactic type S. But here the content of S is, *there is a sphere*. Or is it? Is local supervenience really violated? Behaviourally, P and R are indistinguishable: *e.g.*, both trace a cube-like path. And R will hardly be well adapted to an environment of spheres when he has cube-like experiences; imagine him trying to build a shelter from spheres. It seems the anti-individualist thought-experiment must suppose that the subject does nothing with the cube that depends on its being a cube. But then that removes any motivation for attributing different contents.

The implausibility of this kind of Burgean thought-experiment is that individuation of content is by input only. In a word, his thought-experiment implies a causal covariance theory, that content of a type of experience varies according as the causal antecedent of that type of experience varies, for nowhere does he advert to output. Yet it is *prima facie* implausible to ascribe the same content, on the basis of an invariable distal cause, where the

behaviour of P and R varies as a result of different experiences. That so, the Internalist can accept the second option, for the disjuncts will not be indefinitely long. Hence both contents crack and crackdow can figure in Intentional explanation. Then the content of intrinsic type T experience will be: *shadow in N and crack in W*, but not *crack in N and the shadow in W*. Moreover, the Externalist is mistaken in saying that discriminatory ability can not be made sense of, for he assumes that no cracks can be distinguished from crackdows. But only some cracks are indistinguishable, and therefore only some are crackdows.

Thus the Internalist accepts the pawn, but discards the poison.

**§4.7. A Revised Thought-Experiment.** Be that as it may, nothing debar the Externalist from accepting that, because of environmental changes, perceptual content can be preserved across duplicates. Nothing debar him from accepting that different distal causes may have the same experiential effect. For this alone does not establish Individualism. As it happens, Burge nowhere says that distal cause is a *sufficient* condition for ascription of perceptual content.<sup>10</sup> An alternative option, one that he himself would accept<sup>11</sup>, is to say that evolutionary factors are the key feature in attributing content. Now the refutation of Internalism requires only that *an instance* in which the supposition of strong local supervenience fails.

In which case, by making content answerable to behaviour, the revised thought-experiment will hold the internal constitutions of the duplicates constant, whilst varying perceptual content *and output* with their distal antecedent cause. But even although behaviour is characterized externally, thus making it a function of environmental factors, it retains its basis in the subject's internal constitution.<sup>12</sup> Consider:

---

<sup>10</sup> Ibid. pp. 86-98.

<sup>11</sup> "Individualism and Psychology", p. 40.

<sup>12</sup> M. Davies, "Perceptual Content And Local Supervenience". *Proceedings of the Aristotelian Society*, 1992. §2.3.

In the actual situation, P's intrinsic T-type experience is caused by O. His B-type behaviour, which is consequent on his T-type internal state, is appropriate to occurrences of O. Thus input O and output B are in harmony. Moreover, this harmony accords with P's evolutionary history. In the counterfactual environment, however, R's intrinsic T-type experience is caused by antecedent C. Yet his D-type behaviour, which too is consequent on his T-type internal state, is appropriate to occurrences of C. That is, input C and output D are equally in harmony. Moreover, this harmony accords with R's evolutionary history.

Thus the Externalist can grant that phenomenological character of T-type experience is locally supervenient on T-type internal states and thus preserved across duplicates, *i.e.*, P and R. He accepts too that phenomenology supervenes on perceptual content. The perceptual content of any subject *x* cannot be different unless the phenomenology of his experience is. But this later proposition involves a weaker supervenience claim. For content here supervenes on the intrinsic experience of a subject *x* within the *same possible world*. Whereas the internalist makes the stronger claim: that perceptual content of duplicates *x* and *y* supervene on their phenomenology *across possible worlds*. The first concept relies on counterfactuals within the same logically possible world; the other relies on counterfactuals across them. The Internalist holds that perceptual content supervenes on phenomenology in the sense that for any two individuals, *x* and *y*, and in two possible world, there can be no difference in phenomenology, and therefore perceptual content, without a difference in the subjects internal states. By contrast, Davies holds that, within a logically possible world and within an individual, the perceptual content can differ only if the phenomenology of his experience differs. Thus the Externalist can accept the proposition that "experience" has an intrinsic character, and that this is strongly supervenient. That is, he can accept that for any subject a perceptual content can differ only if the intrinsic phenomenal character of the experience differs. But he will reject that perceptual content across duplicates supervenes on phenomenology; and by transitivity, on internal constitution. That is what makes local supervenience strong, and that

is what the Externalist claims to have refuted. For, on the basis of both input and output, R is ascribed with intentional state type C; and, on the same basis, P is ascribed intentional state type O. Yet their internal constitutions are the same. The duplicates have the same phenomenological character of experience; it supervenes on their internal states. Yet their perceptual contents differ. So however similar their internal states, they have the same intrinsic T-type experiences; and however different their distal antecedent causes, environmental facts ensure, perhaps by way of an optic law, that R's behaviour (D-type) differs from P's (B-type) behaviour, thus showing that strong local supervenience is false.

Though the Externalist appears to have disarmed the Internalist's strategy of motivating the ascription of a common content on the basis of behavioral output, namely by making behaviour accord with input in each case, the Internalist should reject as implausible the suggestion that there is a distinction (at least in the cases considered above) between the perceptual content of an experience and its intrinsic phenomenal character (to say nothing of the bizarre counter-example forced on him by his equally bizarre intuitions about this kind of case). For surely phenomenology of the above experience is nothing if not representational.<sup>13</sup> So it is far from clear that the Externalist has won.

This, however, is far from the end of it. For we have yet to examine a still different kind of Externalist strategy: namely the full-blooded singular thought thesis of Gareth Evans. That examination, however cursory, will suffice to highlight the difference between Externalism *vis-a-vis* general and singular thoughts. It will also help us to throw light on a problem for Burge, a problem which I shall exploit on the Internalist's behalf.

---

<sup>13</sup> See Peacocke, ch. 1. (1983), on *sensational* properties of experience. For criticism see also McGinn, 1989, p.73-75, and Dennett 1988 in Lycan (ed) *Mind and Cognition*.

## Chapter Five

---

**§5. Semantic Internalism and Externalism: A Synthesis.** We saw in chapter 1 that the question of external individuation of both semantic and Intentional properties can be motivated by asking the broader, though connected, question of how the mind manages to refer to external objects, the idea being that to know this is to know what is involved in the individuation of mental states.

One suggestion was that a causal relation obtains between an expression's meaning (or propositional attitude) and the object in the world. So for example, with respect to the semantics of a proper name, a child might be introduced to "Socrates" having heard only that its bearer was a Greek Philosopher,<sup>1</sup> and still, thanks to the appropriate causal relation, be counted as referring (the appropriate link typically being, according to Kripke<sup>2</sup>, one which passes from those present at the referent's naming baptism through to future users of that name).

On the other hand, it is said that insofar as this causal link can, and often does, obtain without impinging on the subject's consciousness, the above account of Intentionality (*i.e.* of how the Intentionality of names can be extended to apply to Intentionality generally) falls short of the mark, for it says nothing of the fact that the relation we wish to explicate (between the mind and the world) is essentially conceptual, and hence that by ignoring this aspect of our cognitive lives, the above kind of Externalism dispenses with the surely important idea that the user of the name must, if he is to succeed in referring, have a conception of how to distinguish Socrates from other objects in the world. Thus since content is constituted rather by the subject's conception of the thing referred to, the nature of reference is partly explained in terms of one's possession of a concept or Fregean Sense, in particular, by the fact that reference is achieved where something satisfies the concept associated with the term.

But then we saw that insofar as oneself and one's doppelganger are said

---

<sup>1</sup> Cf. Kripke. See Evans, *The Varieties of Reference*, p. 74 (n.13).

<sup>2</sup> See his *Naming and Necessity*: Oxford.

to share the same psychological state, they might be said to share (in some sense) the same concept despite the fact that, situated in different environments, they refer to different things,<sup>3</sup> thereby casting the above explanation of reference into doubt.

An attempt to break the stalemate (apart from our own) - between on the one hand the view that reference is purely a matter of a conceptual relation, and on the other that it is purely causal - has been made by maintaining that whereas the causal relation is a necessary condition for reference and representation, it is not sufficient, since a conceptual element, mediating between the mind and the world, must also figure as an explanatory fact. So for example, the introduction to "Socrates" requires not just the kind of causal relation outlined above, but also an appropriate conception of its bearer, where again that conception amounts to a piece of discriminatory knowledge by means of which the subject can pick out determinately the referent of the name, and where that conception itself is governed by the appropriate causal link. This is the view propounded by Evans. Here in more detail is what he says.

**§5.1. A Theory of Singular Thought.** Evans takes Frege's metaphor "mode of presentation" as containing the observation that when we think of an object we do so in one of potentially many particular ways, such that since an object is what a proper name refers to, a *grasp* of the name's *sense* involves a *capacity* to think of its bearer in such and such a way. This suggests to Evans that to have that capacity is to have a route to identifying the object; a route, moreover, that is constituted by one's having individuative or discriminatory knowledge. As we will see, this is a special way of saying that the sense of a proper name helps determine the truth-conditions of the sentence in which it occurs; and also that it is a constituent of the proposition entertained. Evans further claims that the constituent concepts of a singular thought are tied to concrete objects, on the grounds that being able to think of an object in a

---

<sup>3</sup> See §1. above.

certain way rests on there being an *identifiable* such object to think about. In other words, the act of judgement involved in the predication of a property to an object is itself tied to the thing in the world, for the reason that the identity of the object figures in the identity of that judgement, and *a fortiori* to the identity of the thought. Thus Evans holds that with singular propositions the object must exist for the judgement to be had, and moreover, that for the judgement to be particular - for the singular thought to have particularity - the object in question must be singled-out determinately, since only by identifying it can we grasp the sense of a proper name as it figures in the singular proposition, and only thereby can we successfully refer.

To put the point in a still greater Fregean light, Evans says that a proper name introduces an object, its referent, as an argument to a function which, introduced by the concept-expression (say) " $\partial$  is wise", then maps to the value True all objects falling under that concept-expression " $\partial$  is wise". But an argument can be presented to the function *only* if a singular term has an object as its reference, for only on having a truth-value - be the proposition asserted or not - can the sentence have significance, for the reason that the semantic power of a constituent expression, whose value derives from association with extra-linguistic entities, affects the truth value of the sentence it helps form.<sup>4</sup> To further the linguistic analogy:

"... a term is a Russellian singular term if and only if it is a member of a category of singular terms such that nothing is said by someone who utters a sentence containing such a term unless the term has a referent - if the term is empty, no move has been made in the "language-game". To say that nothing has been said in a particular utterance is, quite generally, to say that nothing constitutes understanding the utterance."<sup>5</sup>

Thus without a truth-value, without a referent or object in the world, no determinate singular thought or proposition is expressed or entertained. Again because the constituent sense which contributes to the formation of the thought is tied to the object. By corollary, we can say that a belief cannot be ascribed

---

<sup>4</sup> Similarly with concept-expressions. See Evans, *ibid.* p. 11.

<sup>5</sup> *Op. cit.* p. 71.

to a speaker who [purportedly] understands a sentence containing an empty proper name. For if the sentence lacks a truth-value, the belief will be neither correct nor incorrect.

What, now, of discriminatory knowledge? Note from the above that the subject knows the truth-conditions of an atomic sentence only where he grasps its sense: *i.e.* the *thought*. Note further that the grasp of sense, the grasp of a determinate thought, is a conceptual skill the possession of which is manifest in the possession and exercise of discriminatory knowledge: this because to have grasped the requisite concepts is to have grasped the truth-conditions of the proposition it helps constitute. In fact the grasp of a singular thought requires the possession of two kinds of conceptual ability: that skill equated by Evans with the possession of a concept of an object, and that of a concept of a property. So our conceptual skill comes into play precisely in singling out the object determinately. Now this is partly a matter of one's discerning what Evans calls the object's "fundamental ground of difference":<sup>6</sup> facts which distinguish the object (at a particular time, for material objects) from all others: e.g., a shape is differentiated by its geometrical properties, a shade of colour by its phenomenal properties. And because *ipso facto* every object has a fundamental ground of difference, and because one's conceptual ability to discern it is a constituent of one's singular proposition, we can see still more clearly why Evans says that the thought itself is dependent on the object, and can also see more fully what he means by saying that identifying the bearer of a proper name is to think of it in some particular way, and that what we know when we understand a name is the truth-conditions of the sentence in which the name occurs.

Putting these points together, we can say that to entertain a singular proposition about a concrete object is to know what it is for a proposition of the form [a is F] to be true, where that knowledge lies in virtue of the truth of a pair of propositions, [ $\partial = a$ ], and [ $\partial$  is F], where  $\partial$  stands for a fundamental idea of an object, such that the grasp of each constitutes, respectively,

---

<sup>6</sup> Op. cit. p. 107.

knowledge and therefore possession of both a concept of an object and a concept of a property.<sup>7</sup> Similarly, for a time relative concept of an object  $\partial_t$  and time relative concept of a property  $F_t$ . Possession of the concept of a property  $F_t$  is a matter of knowing what it is for a proposition of the form  $[\partial_t \text{ is } F_t]$  to be true. Discriminatory knowledge is thus also about tracking the object in space and time. With the proposition  $[\partial_{t_*} \text{ is } F_t]$ , where  $t_* \neq t$ , knowledge of the identity conditions of an object through time, and knowledge of what it is for an object to possess a property, is a matter of the joint truth of a pair of propositions,  $[\partial_t = \partial_{t_*}]$  and  $[\partial_t \text{ is } F_t]$ , where knowledge of the truth-conditions of the first of the pair of propositions is what it is to possess a concept of an object in time, whereas the latter is knowledge of the possession of a concept of a property through time.<sup>8</sup>

We can now say, on Evans behalf, that the importance of possessing discriminatory knowledge lies in the importance of the subject in being able to *identify* the object of his putative thought, for without which ability he cannot know what particular object he purports to be thinking of, and so cannot thereby have grasped the constituents of the proposition he wished to entertain. Hence unless he knows which object his putative judgement rests upon - unless he knows what he judges - he cannot have grasped a singular thought.<sup>9</sup>

**§5.2. Demonstrative thoughts.**<sup>10</sup> Let us return to the earlier point that the causal link and mode of identification are together necessary and sufficient for the subject to refer, and let us combine this point with what we've just said about the latter of our dual components. Then according to Evans, what we might take to be a demonstrative proposition about a person we see on (say)

---

<sup>7</sup> Op. cit. p. 111.

<sup>8</sup> Op. cit. p. 111.

<sup>9</sup> Op. cit. pp. 92, 139-40, 170, 173.

<sup>10</sup> I have space for a discussion only of demonstratives, these being the "mother and father of all such information based" singular thoughts, op. cit. p. 145; what follows therefore precludes discussion of recognitional and self identification, see his ch. 7 & 8.

the television, like "That man is F", is in fact no such thing. It is true that we are disposed to treat the object as relevant to the truth-value of our proposition, that is we know what it would be for a proposition of the form [a = that man] to be true.<sup>11</sup> But the existence only of the causal link means that the kind of thought involved has not the conceptual simplicity of a genuine demonstrative thought. Here the truth of the identity proposition [ $\partial$  = that man] is such that the object or man is thought about only as being causally responsible for the sounds and images perceived, where this additional conceptual element involved in discriminative knowledge consists of the idea of tracing the immediate object of perception back to its causal source. Yet this may not enable us to locate the object. By itself the causal link "cannot provide the subject with a knowledge of what makes it the case that an object, distinguished as the occupant of a position of space, is that object".<sup>12</sup> Then the kind of concept possessed in virtue of this information is such that the content of thought expressed by "That man is F" is general or descriptive rather than singular.

This type of [descriptive] content is not, we are told, always available where the appropriate components fail to conjoin. Evans has us imagine a savage confronted with a radio for the first time who exclaims "that man in the box is very small".<sup>13</sup> Because he altogether lacks a conception of perceiving someone distant in space and (perhaps) time, he lacks not only a concept of "that man" but also the concept of a man causally responsible for those sounds, and so cannot be ascribed any such thoughts at all, demonstrative or otherwise. And without the demonstrative concept of the man, he's unable to identify the object of his putative thought, and is therefore unable to specify the content of what he means by saying "that man" whom he hears. Again the savage must know what it would be for a proposition of the form  $\delta = a$  to be true, as he cannot think about what he cannot identify.

---

<sup>11</sup> Op. cit. p. 147.

<sup>12</sup> Op. cit. p. 149.

<sup>13</sup> Op. cit. p. 149-50.

We can help our understanding of what it is to locate an object, suggests Evans, by understanding what it is to have thoughts about places. The identification of places involves a frame of reference, where places are distinguished in terms of their relations to the objects which occupy them. This frame of reference amounts to the possession and exercise of a "cognitive map": notably our ability simultaneously to represent spatial relations among distinct things.<sup>14</sup> Now according to Evans, our possession of this ability, and with it our ability to have demonstrative thoughts, tells against Cartesianism. For it is apparently because we can form and employ representations of this kind, that we possess a genuine concept of an objective world; and it is because we have a genuine concept of objects existing in space independently of ourselves, whose identification with our cognitive map allows us to express such thoughts in other than such egocentric terms as (say) "here" and "there", that our thinking about the world at this basic level is objective.<sup>15</sup>

Evans brings out his anti-Cartesianism by asking how it is possible to have egocentric spatial thoughts. It is true that such a capacity rests partly on our having a conception of egocentric space; but the very possibility of our having a conception of egocentric space and thus the possibility of our having egocentric thoughts, itself rests, he says, on the fact that we have dispositions to act. And it is apparently because we have such dispositions that we are able to effect a coincidence between private and public space, and thus are able to acquire a genuine concept of objective space. For unless we have the capacity to forge the connections between perceptual inputs and action, which is another way of saying what he means by effecting a coincidence between private and public space,<sup>16</sup> we cannot make the above identifications about places. In which case the perceptual inputs would lack any spatial significance to us. So

---

<sup>14</sup> Op. cit. p. 151.

<sup>15</sup> Op cit. p. 152.

<sup>16</sup> Which is also to say that Evans doesn't mean that private and public space are different kinds of space, just that they're different kinds of information about space: though the former, unlike the latter, is specifiable in egocentric spatial terms, such terms still refer to points in a public three-dimensional space. See op. cit. p. 157.

if we can reason and act such that (say):

"If I am here, midway between Balliol and the Bodleian, then that must be Trinity, and so the High must be down there."<sup>17</sup>

then through this possession and exercise (by action) of our "cognitive map" through action, we can effect the coincidence between egocentric and objective space. And it is just this ability to find where in the world one is that warrants my being credited with discriminatory knowledge about places - that is knowledge of what it would be for identity propositions of the form [ $\pi = p$ ] to be true (where  $\pi$  stands for our holistic identification of a place, *i.e.* use of our cognitive map) - and thus warrants my being credited with demonstrative thoughts about positions in space. And according to Evans, only where this is possible can we say that the subject's egocentric space is space at all (where such ascriptions themselves require of us to think of the subject's situation from an objective point of view).<sup>18</sup> Evans makes a similar point with respect to our discriminations of the direction of a sound:<sup>19</sup> namely that we have no perception of its direction until its direction can be specified in egocentric terms, where the meaning of such terms expressing this relation to one's body - (say) "It's F over there", "It's F up there to the left", "It's F a bit behind me", ("here"-thoughts being the least specific of this species) - are themselves determined by our dispositions to act.

Having explicated the nature of egocentric-spatial-thoughts, Evans goes on to stress the differences between these and demonstrative thoughts proper. He considers a man who lying in the dark having placed a bottle of whisky beside his bed, says "there is a bottle of whisky there". That he has the capacity to select a definite place in egocentric-space, and thus has a definite thought, even although he is not perceiving that place, is manifest in his maintaining a stable dispositional connection with the place he has in mind: he knows in other words what it would be for an identity propositions of the

---

<sup>17</sup> Op cit. p. 162.

<sup>18</sup> Op. cit. p.163.

<sup>19</sup> Op. cit. p. 155-60.

form  $[\pi = p]$  to be true. As Evans puts it:

"If the subject does have an Idea of a place (does know which place his thought concerns), this will be manifestable only in manifestations of that stable dispositional connection: thus, in his treating certain perceptions from that place as unmediately germane to the evaluation and appreciation of any thought involving the Idea, and in his directing actions towards that place when thoughts involving the Idea, together with other circumstances, indicate that this is a good thing to do."<sup>20</sup>

By contrast, you might be in the dark, struck with the thought that there is something immediately in front of your nose, and so might be disposed to act in the way you would were the lights on.<sup>21</sup> That the thought involved would not be demonstrative is because unless one is perceiving an object one cannot in general know that one is dispositionally related to an object. Whereas again:

"Places ... being - how shall we say? - so much thicker on the ground than objects, a subject cannot fail to have a single place as the target of his "here"-dispositions at an instant ..."<sup>22</sup>

This distinction has Evans consider a case where a television screen shows pictures of a seabed sent by a remote-control submarine.<sup>23</sup> It would be natural to suppose (as we did *mutatis mutandis* with the savage-and-the-radio) that a subject could think, whilst watching the pictures on the mainland, that *It's murky here*, etc. If he had a conception of the situation - as the savage in the analogous case did not - the content would be of the place as where the submarine is, *i.e.* descriptive. On the other hand, were he unaware that they are television-images, and were he convinced that beyond the glass lies the bottom of the sea, then his attempt at an identification of the place in egocentric space would fail. As Evans says:

"... we would not need to say that his thoughts concerned this

---

<sup>20</sup> Op. cit. p. 161.

<sup>21</sup> Op. cit. p. 169.

<sup>22</sup> Op. cit. p. 169.

<sup>23</sup> Op. cit. p. 164.

place: this is a case of a thought which is not well-grounded".<sup>24</sup>

Which, I take it, is to say that:

"If there is no place thought about, there is no thought at all - no intelligible proposition will have been entertained".<sup>25</sup>

So provided you're not, unbeknownst to yourself, hurtling through space - to cite another case<sup>26</sup> - in other words, provided you can keep track of the place you occupy, so that your maintaining a stable disposition with it allows you to treat the place as germane to the truth of a "here"-thought, such a proposition can be entertained.<sup>27</sup>

Let us now see what is wrong with the above. By doing so, we will be ready to provide a further defence of Internalism.

---

<sup>24</sup> Op. cit. p. 165.

<sup>25</sup> Op. cit. p. 170.

<sup>26</sup> Op. cit. p. 170.

<sup>27</sup> Op. cit. p. 164. p. 170.

## Chapter 6

---

**§6. The Reducibility of Singular Propositions.** It is not uncommon as a first reaction to feel an ambivalence towards Evans' theory of singular thought: to believe that singular reference is on the one hand intimately related to the object referred to, and to believe on the other hand that it cannot be - on the grounds that when I speak of a particular object, say of a painting that has impinged itself on my consciousness, it is not to that particular object in world ( $w_a$ ) that I must refer in order to identify the object of my thought, for by bearing in mind that an indistinguishable experience of that kind could be caused by another object in world ( $w_r$ ), say a replica picture by an equally competent painter, or by no object at all in world ( $w_h$ ), then it seems that, *pace* Evans, the importance of the referential relation, in identifying the thought, lies not with the object in the world, but with the experience itself.

Simon Blackburn pushes this latter line.<sup>1</sup> It is true, he would say, that the demonstrative used in  $w_a$  and  $w_r$  succeeds in referring to an object whereas none is referred to at all in  $w_h$ ; and true too that an intention to refer to a particular object suggests that the thought be described in terms of states of affairs in the world (*i.e.* truth or falsity of what is predicated of the object). Nevertheless the identity of the thought itself, he asserts, is given not by reference to any object, but by reference to the *modes of presentation* (the "experience"). And because from the first-person perspective the experience is constant across the three possible worlds ( $w_a$ ,  $w_r$ ,  $w_h$ ), so too is the thought. Indeed, says Blackburn, it is not just that the first-person perspective is the same; the proximate input from the environment - such features as sound and light waves (including Cartesian scenarios) - are also constant. Thus though an apparent singular thought may have been *expressed*, the thought *actually had* is universal or descriptive. There is then no self-standing category of singular thought.

Searle holds a similar line about singular thoughts.<sup>2</sup> It is true, he says,

---

<sup>1</sup> Simon Blackburn, *Spreading the Word*, chapter 9.

<sup>2</sup> John Searle, *Intentionality*, esp. Chapters 8 & 9.

that (i) there are *de re* beliefs in the sense that they purportedly refer to actual objects, and that (ii) they can't be described in general terms. But they're nonetheless a subclass of *de dicto* thoughts and beliefs. Furthermore, this latter class, he maintains, are independent of the world: a Cartesian subject could have them. Searle backs up this Internalism with a theory of perception, to which these two propositions are crucial:

- (A) *that visual experience is wholly propositional and,*  
 (B) *that its Intentional content is "self-referential".*

The visual experience is Intentional because it is directed at something; that it is propositional is because it has truth-conditions; and that it has truth-conditions is because the visual-experience is implicitly "self-referential": it specifies that the *experience* be caused by the rest of its conditions of satisfaction, *i.e.* by the state of affairs, if any, perceived. So from the experience alone the subject knows what conditions in the world must obtain for it to be a perception rather than (say) an hallucination, or for one's belief to be true or false. So for example:

*I have a visual experience that there is a yellow station wagon there and that there is a yellow station wagon there is causing this visual experience.*

Specifically, the conditions are: (i) that there be a yellow station wagon, and (ii) that the Intentional object be the cause of the visual experience, where (ii) specifies a criterion - a causal link - for being a perception rather than an hallucination:

Clearly, the "self-referential" component of Intentional content is crucial as without it no Intentional object could be individuated solely by the Intentional, purely conceptual, content. So what makes *de re* a species of *de dicto* is that because Intentionality is intrinsic to the visual experience, and because the visual experience is self-referential, it already contains content with determinate truth-conditions. Moreover, that Intentionality is intrinsic rests on the premise that the nature of Intentionality is a first-personal subjective phenomenon, a matter of phenomenology. Thus Searle assumes that

the same Intentional state could be had regardless of the cause:<sup>3</sup> the contents of a Cartesian subject would be fully truth-conditional, only all his thoughts would be false.

Note that here Searle's animus is explicitly against Burge rather than Evans; I have simply transposed the thrust of Searle's contention - *i.e.* the collapsing singular thoughts to general thoughts - to Evans.

We can now say that though both Evans and Burge (1977) hold that singular thoughts are *sui generis*, the difference between them is that where for Evans they are fully conceptualized, whose content is truth-conditional, for Burge the content is expressed by an open sentence, whose subject position is occupied by a free variable, and whose referent is determined by the non-conceptual element of the context. As already seen, to Burge the conceptual element, the predicative fragment, is where Externalism lies, as this element characterizes the mental state. Thus though it is true that Searle's (and Blackburn's) thesis concerning the primacy of *de dicto* would undermine Evans' Externalism, a further argument is required to counter the Burgean claim that the individuation conditions of *de dicto* contents are External. I provide this further argument in §6.4. below.

Lastly, it is interesting to speculate whether Burge's move from his weak singular thought Externalism of 1977 to his strong general thought Externalism of 1979 was made not in fact just because he saw that his early theory was too weak to be interesting, but because to make for a strong singular thought Externalism would require of him to admit of their being fully conceptual, and that that, with illusory thoughts in mind, would be to go too far. If that speculation is right, then as we will see below the move of 1979 was misplaced.

**§6.1. Tempered Singular Thoughts.** Peter Carruthers takes a different line to Blackburn and Searle: he grants that singular thoughts are irreducible to

---

<sup>3</sup> Op cit. p. 213.

"universal thoughts", but denies that sense is object-dependent;<sup>4</sup> the content of one's demonstrative thoughts, he argues, is constant across the above substitute possibilities. This is not to say, however, that their truth-conditions, and thus truth-values, are the same. According to this view, truth-value belongs with thought-tokens rather than to the thought-type (the content). So for example, suppose that my doppelganger and I are lion catchers and that I confront a real lion, whereas my twin hallucinates one. Nonetheless, we both grasp the same content-type expressed by "that the lion before me is worth catching"; according to Carruthers, it is just that our token assertions differ: mine but not his has a truth-value.<sup>5</sup> In a not too dissimilar way to the private concept strategy discussed in chapter 3, he views truth-conditions as determined by us, by our (often unconscious) capacities to classify things regularly over time, and to locate them in space: that my particular thought concerns a particular object is because, in being able to find my way round an objective spatial world, and in our being able to locate objects within it, I'm able to project my thought-content onto the individual in question.<sup>6</sup> And thus it is because myself and doppelganger differ in our capacities to find our way round our objective spatial worlds, and in being able to locate the objects within it, that generally speaking our thought tokens have different truth-conditions; but far from that being a reason to think that those thoughts-types are distinct, we have, on the contrary, reason to suppose that they are the same: namely that because we're doppelgangers our experiences and beliefs when construed narrowly (*i.e.* independently of reference) are indistinct.

As my remarks on the private concept strategy of chapter four indicate, I'm sympathetic to the spirit of Carruther's approach. But why say that the proposition grasped in the empty possibility is singular? He says that where the thinker hallucinates, and is known by others to be doing so, the content must be described indirectly; after all, we can't specify the thought with a

---

<sup>4</sup> "Russellian Thoughts", *Mind*, 1987.

<sup>5</sup> *Op. cit.* p. 30.

<sup>6</sup> *Op. cit.* p. 32.

"that"-clause where no such object exists. But that, he believes, hardly counts against treating such thoughts genuinely demonstrative. True it is demonstrative; but we can still hold that the content be reduced *al la* Searle, and not least because without a truth-value it is unclear what an asserted singular proposition could be, and thus it is unclear that it would be singular at all.

To be sure, Carruthers asks how it could be anything but irreducibly *de re*.<sup>7</sup> Experiences, he says, are "transparent" and "unnoticed", and to which no attention is normally paid. Then because one doesn't think of the experience, one's demonstrative thought can't have the same content as a descriptive proposition. Again I'm unconvinced: for one might say, on the contrary, that experiences cannot but be noticed, this because one can't stand outside of one's own experience, experience and object being in this sense pretty much the same. This would answer his second argument, which is that thoughts about particulars are logically prior to thoughts about experiences, such that the latter are possible only because the former have been located in space outside oneself: "how", he continues, "could the one thought content be logically prior to the other if they are, on the contrary, the same content?"<sup>8</sup> But this assumes with Evans that demonstrative thoughts are possible only because we can locate objects in outer space. Our being able to have the same thought in empty possibilities suggests otherwise. In short: it begs the question against Methodological Solipsism.

Harold Noonan also grants that demonstrative thoughts are *sui generis*,<sup>9</sup> and that singular thoughts are made possible as much by our capacities and the dispositions underlying them, as by the world. Moreover, he grants that because, in Evan's savage-and-the-radio-example, the savage

---

<sup>7</sup> Ibid, p. 22. n.8.

<sup>8</sup> Op. cit. p. 24.

<sup>9</sup> "Russellian thoughts and methodological solipsism", in *Language, Mind, and Logic*, ed. J. Butterfield. p. 76-77.

lacks the appropriate mode of identification,<sup>10</sup> the savage is not entertaining a demonstrative thought about the man at the other end of the informational link: he "has no idea what it could be for someone to be *that man* (whom he hears)".<sup>11</sup> But it's quite another matter, says Noonan, to hold that the savage expresses no demonstrative thought at all: quite another matter, in other words, to say that modes of identification are Russellian. Here he quite rightly insists that there is *something* in common - namely the capacity to have a demonstrative thought - not only between there being a little man in the radio and there not being one, but also with there being a man inside who is silent. This, he says, is because (i) the dispositions are unchanged, (ii) because we should equate dispositions with a mode of identification of an object, and (iii) the thought itself is determined by what one is disposed to do (so that since the dispositions are unchanged so too is the content).<sup>12</sup> So why, asks Noonan, suppose that one can't think of an object to which one is dispositionally connected but from which no information is received? Why assume that all demonstrative thoughts must be *about* particular objects? Elaborating on the television-example, he has us imagine a screen showing pictures of the bottom of the sea which covers the wall of a room such that the subject takes it to be the side of a huge aquarium; coincidentally there is an aquarium on the other side. Suppose the man thinks, "there's a large rock over there". Such descriptive egocentric spatial thoughts would be allowed by Evans: the dispositional connection with the location is maintained). Now Noonan rightly argues that it would be arbitrary to rule out, as Evans is committed to do, the demonstrative thought "that rock over there is a large one". True, we can't say that the demonstrative thought *concerns* a rock in the aquarium, the coincidence is irrelevant, though we can say that the former thought is about the location in the aquarium, this being a consequence, again not of the coincidence, but of the fact that places are "so much thicker on the

---

<sup>10</sup> Op. cit. p. 87.

<sup>11</sup> Op. cit. p. 84.

<sup>12</sup> Op cit. p. 82.

ground". As just seen, the coincidence shows that the presence of the object is irrelevant to the presence of an undifferentiated stable disposition. Yet it is just this kind of coincidence which Evans would rely on in order to distinguish the above three possibilities of the savage and the radio.<sup>13</sup> Whereas what matters for the existence of any demonstrative thought is that a stable disposition be maintained. Noonan then suggests that where the demonstrative thought seems to be about an object, they will "be dispositions, concerning a certain place, to direct actions to whatever object of a certain type is located there ..." <sup>14</sup> and hence that these dispositions determine that the thought in question is *about* the place which one takes the object to occupy.

It should be clear from what I have already maintained why I think that Noonan is right to say that dispositions are partly constitutive of content. It should also be clear why I think Noonan mistaken in supposing that the demonstrative thought in question concerns a place rather than an object. The source of doubt is that the subject intends to think about an object rather than a location. This can be supported by my defence of Searle; but also by something which Noonan himself says: namely that the subject maintains a stable disposition because of the sameness of how things seem. Of this more later.

**§6.2. Psychological Explanation.** All the above anti-Russellians, with the exception of Searle, reject any suggestion of Methodological Solipsism: they deny that an account of thought-content (and mental states generally) can dispense with reference to facts about the physical world; for they deny that the very same states can be had by a disembodied mind. Rather their fire is towards Russellianism: true, singular thoughts about physical objects in the world require the existence of a world containing such objects; but these thoughts, they assert, can be described as distinct from each other without referring to a particular object.

---

<sup>13</sup> Op. cit. p. 88.

<sup>14</sup> Op. cit. p. 88.

To this end, they offer their [alternative] accounts on the grounds that the notion of "illusory thoughts" is theoretically unmotivated. For suppose that I perceive a lion whereas my doppelganger hallucinates one and that being lion catchers we attempt to catch the one we think we see. How might we explain this action? The Internalist notes the inductive plausibility of holding that different causes can give rise, as suggested here, to the same experience or effect, inferring thereby that both actions are of the same kind. Hence if the experience and action is the same, we have every motivation for ascribing the same propositional attitudes. In fact, the above anti-Russellians strategy works in two-stages: first, where there are behavioral and experiential symmetries we ascribe the same thought-content to myself and my doppelganger; second, mention of facts is then made to indicate that the beliefs were true or false (or to make clear that they're instances of knowledge).

Indeed Carruthers offers a diagnoses of the Russellian's mistake: namely its failure to distinguish between two purposes in content ascription. Specifying content with a "that"-clause is important, he says, where the purpose is belief-acquisition. Clearly, a specification of the precise object can prove important in learning what someone believes. But no such clause is required, he says, where the purpose is to explain behaviour.<sup>15</sup> So for example, though the capturing of the lion can be interpreted with reference to the object itself (that *that* lion was thought by me to be worth capturing) the content of this *de re* ascription is related non-essentially to the object, for the truth of the relation between my thought about the object coincides with the truth that I have (to echo Blackburn) a "universal" lion-possessing thought; it is simply caused by and focused on that particular object, but could have occurred had another object of that kind (or no object at all) entered the fray.

Hence we iron-out the [external] asymmetries (*i.e.* the fact that each act was directed at different objects) and hold that since the actions are thereby the same, so too are the thoughts which caused them: for example, we replace in our explicanda "X ρ-ed Y" with something like "X ρ-ed", where ρ-ing is a

---

<sup>15</sup> Cf. Blackburn, *op. cit.* pp. 325-327. Also Carruthers, *op. cit.* pp. 26-28.

common movement, like reaching out with a lion net. So since the phenomena in question falls under the same description D, we can expect, *ceteris paribus*, an explanatory description D\* applicable to our above cases, and can say of our subjects that (e.g.) each *seemed to be confronted by a lion worth catching*.

Hence the idea of illusory singular thoughts might seem unmotivated: as indeed might the general [Russellian] idea that psychological states differ according as the identity and existence of objects differ, where this more general idea is partly the thought that different explanations must be provided where different objects are concerned, even where those objects are of the same kind. So where my doppelganger but not myself hallucinates a lion, the Russellian is committed to provide different Intentional explanations for what appears to be exactly similar behaviour, just as he is *mutatis mutandis* where we perceive distinct lions.

Insofar as we abstract the subject's environments, you might wonder, however, just how anti-Methodological Solipsism this anti-Russellianism is. As McCulloch<sup>16</sup> points out, whereas to iron-out the asymmetries from the Intentional descriptions is to be close to ignoring one's professed anti-Methodological Solipsism (MS), to include them is to come close to the Russellian.<sup>17</sup> Moreover, he notes that the ascription of these symmetries goes against the normal practice of psychology. Why, in any case should we see ourselves and each other in such stark terms: "Why", he asks, "should we become such monsters? Civilization as we know and love it would crumble!"<sup>18</sup> More seriously, inclusion of the environment in Intentional descriptions may be important: for, against Carruther's diagnostic distinction, we may (i) lack the right sort of knowledge of each other to proceed in any other way; and (ii) the object-involving Intentional descriptions may provide a useful source of information about people.

---

<sup>16</sup> G. McCulloch, "Faith, Hope And Charity: Russellian Thoughts Defended", *Analysis* 1988, p. 88.

<sup>17</sup> P. Carruthers *Russellian Thoughts Attacked*, *Analysis* 1988, p. 93.

<sup>18</sup> McCulloch. *op. cit.* p. 90.

You might reply that there are still purely explanatory contexts in which our purpose in descriptions of thoughts isn't the acquisition of new information about the world, and in which there's no lack of knowledge concerning the subject's mode of thinking. So for example, there may be two lions against whom myself and my doppelganger are confronted, and in response to which we both think, *that lion is magnificent*. Do we really want to ascribe different types of thought, as the Russellian urges? At any rate, our common explanatory descriptions need not be so stark. Suppose that on seeing a lion lurking in your neighbourhood, and such that your doppelganger hallucinates one in his, you both think: "that cat is dangerous", "I'd better warn my neighbours", "I'll get my student to catch it". Our Intentional descriptions would quite rich.<sup>19</sup> Again: why ascribe different thoughts (*i.e.* the original singular thought to you, and a descriptive (non-singular) belief to your doppelganger)? The divergence of explanation seems unmotivated when exactly similar behaviour is involved.

Looked at another way, perhaps the behaviours aren't the same: you and your doppelganger live in different neighbourhoods and direct your acts to different things. And anyway, even where action can be subsumed under the same description, different explanations might be warranted. McCulloch again:

"at some time both Carruthers and I set out to become professional philosophers. But it does not seem at all plausible to claim that there should be some parallel folk-psychological explanation for this (we are both perverse individuals?)".<sup>20</sup>

Moreover, the Russellian can reiterate a point made earlier: that in hallucinatory cases, no object is identified, so no content is specified. Lastly, he can argue that his concern is with the preconditions for the possibility of singular thought, not with the explanation of behaviour.

To be sure, one's intuitions with respect to psychological explanation are pulled now one way, now another, according as we vary the scenarios

---

<sup>19</sup> Cf. Carruthers, *op. cit.* p. 95.

<sup>20</sup> "Carruthers Repulsed" *Analysis*, p. 99.

congenial to our purposes. Thus from what we have said so far, no-one gets the last word. However, let me end with a remark just made: that we let our concern lie with the preconditions for the possibility of thought. Then because we have a plausible alternative, we have no compelling reason to accept the Russellian analysis of those preconditions.

**§6.3. Illusory Propositional Fragments.** I promised to provide an argument to strengthen the Blackburn/Searle thesis: for while arguing that singular thought is reducible to descriptive contents, they left unaddressed the possible Burgean retort that *de dicto* propositions are externally individuated.

The philosopher who is sympathetic to Burgean style Externalism may have rejected the Evansian-variety on the grounds that from the first-person perspective the notion of "illusory singular propositions" lacks intuitive appeal, that if anyone can decide on whether or not a singular thought is being entertained it surely is only the subject himself. (Indeed I speculated in §5.5. that in working out his Externalism Burge himself may have been similarly motivated.) That sentiment, I shall argue, is mistaken, as it is far from clear that Burgean Externalism itself precludes the possibility of illusory thoughts. This is not of course to say that no principled choice can be made between Evansian and Burgean varieties of Externalism: after all, one may not subscribe to the demand for discriminatory knowledge.

It is true that the Burgean conception of singular thought excludes object-dependent sense. Recall that for Burge (1977) the content of a singular thought is given by an open sentence, whose subject position is occupied by a free variable, and whose referent is determined partly by the context, rather than entirely by the concepts. Then empty and non-empty singular thoughts can be treated as having the same content, for any object in context is part not of the content, but simply what the thought happens to be about. But now the question arises about whether or not, with respect to general thoughts, Burgean Externalism is property-dependent in a way that Russellianism is object-dependent. And it would seem that it is, since for Burge the properties in the world are those in terms of which the [general] thoughts get their content. The

question gives rise to a hitherto undiscussed notion of illusory general thoughts concerning the putative contents of the world. For it is logically possible that the type of thing said to individuate the predicative fragment is non-existent. Imagine that the world is such that water is unreal, that we are undergoing prolonged hallucination, where the phenomenological qualities (embracing of course all our investigations into its microstructure) are just as if caused by the property itself. *Ex hypothesi* it appears that the predicate would lack content.

This is just the sort of question the full-blooded Internalist must press. So how would the concept-expression get its content? The answer, I take it, suggests itself: content derives from the phenomenology of the visual experience, and is individuated by the internal facts, by our dispositions both to classify things regularly over time, and to act on the grounds of how things seem. So in possessing this conceptual capacity to classify regularly over time - *even where what we classify is unreal, and even where, as hoodwinked Cartesians, it only appears that we are acting* - the predicate might be said to acquire a meaning in virtue of the systematic use of those expressions.

Of course, if the Burgean agrees that the predicative fragment would acquire meaning in this way, then he must say in what sense, if any, the meaning of the predicative fragment would change where, for example, a corresponding property subsequently came into being. It is difficult to see how it would differ. One suggestion, *faute de mieux*, is that the Burgean rely on Internalist conditions of individuation outlined above without commitment to the doctrine *tout court*. Then the Externalist will say - must say - that one's mental states are distinct according as the contents of the world are distinct, however similar the thoughts and experiences are in their "pure phenomenological feel". The idea is of course familiar: the difference in mental state is determined by the third-personal perspective: only where myself and doppelganger mirror the same first and third-personal perspectives do we have the same mental states. Thus if the Externalist excepts this option - and the only other alternative, it seems, is to accept illusory thoughts - then he is committed to saying that although one's mental states share the same phenomenology, they are in fact about different things: in  $W_1$  the thought is

about water since there really is a causal connection with it, whereas in  $W_2$  it only seems as if it is. In  $W_2$  the nature of the thought will be somehow indeterminate, about nothing at all, save the hallucination itself, whose content is explicable in terms of some figment of the mind. However implausible this may seem, and however much it needs to be cashed out - if indeed it can be - the Externalist is committed to some such conclusion if he is to deny, as he must, that *fragment of thoughts* are constant across the relevant possible worlds, and if he is to deny that in circumstances where we retain our above mentioned capacity for entertaining propositions we undergo illusory thoughts.

In which case, what can be meant by the suggestion that the thought had in the counterfactual world ( $W_2$ ) is "indeterminate"? Moreover, if thoughts in  $W_2$  are sanctioned in this world-independent way, why can't thoughts had in the actual world be similarly judged? To repeat: the Burgean is not here abandoning the external constraint on the thought had in  $W_2$ ; for even although the thoughts had in the actual and counterfactual world seem the same to the subject, the external condition renders them distinct. But it is just this which must strike the Internalist as counter-intuitive. I am unable, and in any case unwilling, to acknowledge any difference in the thoughts had in  $W_1$ , and  $W_2$ . At any rate, imagine a situation where a scientific community introduces a term for a natural kind of which they have no causal contact but about whose discovery they anticipate to be imminent. *Prima facie* it is implausible to suppose that the concept is any different after the discovery from what it was when on the basis of abstract speculation the concept was first coined and introduced. Science of course has a gift for creating concepts, whatever the true nature of things: besides, whoever said that the concept of phlogiston, with its imaginary extension, is without content?

Generally speaking, Externalism seems to suggest that the nature of the relation between thought and the things in the world those thoughts concern, is a relation of presupposition. That is to say, if one's thoughts are about property  $x$  in the world, then property  $x$  must exist, and that if it turns out that no  $x$  exists, then one will have been mistaken about one's thoughts. But then it seems to me that the Externalist must concede to more than mistakes about

his thoughts. He must hold out to the possibility that all his thoughts about the external world are illusory: if there are no properties in the world, there is no content to his thoughts.

One can hear an Externalist reply: so what? So much the worse for what we think we know. That reply is all very well: but it seems to me that there is mileage yet to be made for the Solipsist view of himself and the world - thus no good reason to reject Cartesianism, no good reason to suppose that our thoughts could not be had independently of how the world might actually be.

## Conclusion.

---

We began with an exposition of Putnam's Semantic Externalism, and showed how Burge extended the strategy to Intentional states generally. In our short history of Burge's Externalism, we suggested that in developing his Externalism Burge may have been guilty of supposing that, Evans apart, a philosophically interesting Externalism was to be had from the *de re/de dicto* distinction. As it happens since he did not embrace Russellianism, and given his impulse towards Externalism, he was left with a relatively uninteresting explication of the nature of the relation between the mind and the world.

In chapter 3 we examined Fodor's "middle way", and suggested that it left something to be desired: namely a clearer account of "narrow" content; though it was not rejected out of hand, it was argued that such a notion did little for a systematic propositional attitude psychology: it did little, if anything, to make psychological explanation Intentional. The root of the problem, I argued, lay in his acceptance of a tempered semantic Externalism. In chapter 4 it was urged therefore that a more plausible answer could be had by rejecting *tout court* the Externalist strategies advertised above, and in the acceptance, though not without modification, of the full-blooded Internalist's truth-conditional notion of content. That is to say, we accepted Crane's private concept strategy, but suggested that (i) the metalinguistic argument was not required, and in any case not convincing, and (ii) that he should embrace the consequent idea of immunity from conceptual error, which we clarified somewhat, showing thereby how plausible it might be.

That left us with a radical Internalism, which was then further radicalized by our endorsement and defence of Searle's contention that all thoughts contents are essentially *de dicto*, that they could be had in the Cartesian scenario of massive hallucination. In defending the idea that such thoughts are individuated by truth-conditions, specified partly by the self-reflexivity of one's visual experience, I suggested that there is no good reason behind insisting, as Carruthers did, that it is the object rather than the experience to which our attention is directed: the two could not be significantly distinguished where truth-conditions are being specified.

Finally, I drew attention to the fact that though Searle's general thought Internalism may have fended off Russellianism, namely by collapsing singular thoughts to the former kind, something needed to be said in wake of the possible Burgean retort that *de dicto* themselves are dependent for their individuation on properties in the world. Recall that we touted in chapter 6 the idea that Burge (1977) resisted Russellianism because of the *prima facie* implausible notion of illusory singular thoughts to which Russellianism gives rise. But we also saw in that chapter that general thoughts too might be susceptible to this phenomenon. The notion of illusory thoughts, however, was rejected as implausible on the grounds that the conceptual capacities in having a thought - viz. the ability to classify regularly over time on the basis of experience and dispositions - are maintained across the counterfactual scenarios. That then was my defence of Searle, driven as it was, not so much by conviction that this is the way things are, but by the impression that Externalism, and those in the middle of the road, do little to unwedge more radical views concerning the nature of the relation between the mind and the world. I conclude then that, given the material examined above, nothing tells against the logical possibility that the mind is independent of the world: our thoughts and experiences might not presuppose how in reality things actually are.

## BIBLIOGRAPHY

---

- Antony, M. V. (1993) "Social Relations and the Individuation of Thought", *Mind*.
- Blackburn, S. (1984) *Spreading the Word*. (Oxford: Clarendon Press)
- Burge, T. (1973) "Reference and Proper Names", *Journal of Philosophy*.
- \_\_\_\_\_. (1978) "Belief and Synonymy", *Journal of Philosophy*.
- \_\_\_\_\_. (1977) "Belief De Re", *Journal of Philosophy*.
- \_\_\_\_\_. (1979) "Individualism and the Mental", *Midwest Studies in Philosophy* vol. IV.
- \_\_\_\_\_. (1982) "Other Bodies", in *Thought and Object*, ed., A. Woodfield.
- \_\_\_\_\_. (1983) "Russell's Problem And Intentional Identity". In *Agent, Language, And The Structure of the World*, ed. James. E. Tombellin.
- \_\_\_\_\_. (1986a) "Cartesian Error and the Objectivity of Perception" in McDowell and Petit (1986).
- \_\_\_\_\_. (1986b) "Individualism and Psychology", *The Philosophical Review*.
- Carruthers, P. (1987) "Russellian Thoughts", *Mind*.
- \_\_\_\_\_. (1988) "Russellian Thoughts Attacked", *Analysis*.
- Crane, T. (1991) "All The Difference In The World", *The Philosophical Quarterly*.
- Davidson, D. (1984) "First Person Authority", *Dialectica*.
- \_\_\_\_\_. (1987) "Knowing One's Own Mind." *Proceedings and Addresses of the American Philosophical Association*.
- Davies, M. (1991) "Individualism and Perceptual Content", *Mind*.
- \_\_\_\_\_. (1992) "Perceptual Content and Local Supervenience", *Proceedings of the Aristotelian Society*
- Dennett, D. (1982) "Beyond Belief", in *Thought and Object*, ed, A. Woodfield (1982).
- Descartes, R. *Philosophical Writings*, ed., E. Anscombe and P. T. Geach. (Van Nostrand Reinhold (UK) Co Ltd 1954)
- Devitt, M. "Meanings Just Ain't in the Head", in *Meaning and Method: Essays in Honour of H. Putnam*, ed. George Boolos (University of Cambridge Press 1990)
- Dummett, M. (1973) *Frege: Philosophy of Language* (London; Duckworth)
- Evans, G. (1982) *The Varieties of Reference* (Oxford: Clarendon Press).

- \_\_\_\_\_. "Understanding Demonstratives", In *Demonstratives*, repr in Yourgrau. (1991)
- Kaplan, D. (1977) "Dthat". repr in *Demonstratives*, ed. Yourgrau (1991).
- Kripke, S.A. (1980) *Naming and Necessity* (Oxford; Blackwell)
- Fodor, J. (1980) "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology", repr in *RePresentations* (1981).
- \_\_\_\_\_. (1980) "Propositional Attitudes" repr in Fordor (1981).
- \_\_\_\_\_. (1985) "Fodor's Guide to Mental Representation" *Mind*.
- \_\_\_\_\_. (1987) *Psychosemantics* (Cambridge, Mass.: MIT Press).
- \_\_\_\_\_. (1991) "A Modal Argument For Narrow Content" *Journal of Philosophy*.
- Frege, G. "Sense and Reference", *Philosophical Writings*, ed, P. Geach & M. Black.
- \_\_\_\_\_. The "Thought", in *Philosophical Logic*, ed, P Strawson.
- \_\_\_\_\_. "Function and Concept", *Philosophical Writings*, ed, P. Geach & M. Black.
- McCullough, G. (1988) "Faith, Hope, and Charity: Russellian Thoughts Defended", *Analysis*.
- \_\_\_\_\_. (1988) "Carruthers Repulsed", *Analysis*.
- \_\_\_\_\_. (1989) *The Name of the Game* (Oxford).
- McDowell, J. (1984) "De Re Senses", *Philosophical Quarterly*.
- \_\_\_\_\_. (1986) "Singular Thought and the Extent of Inner Space", in *Subject, Thought and Context*, ed., McDowell and Petit (1986).
- \_\_\_\_\_. (1977) "Sense and Reference of a Proper Name."
- \_\_\_\_\_. (1982) "Defeasibility, Criteria, and Knowledge." *Proceedings of the British Academy*.
- McGinn, C. (1989) *Mental Content* (Oxford: Blackwell).
- \_\_\_\_\_. (1982) "The Structure of Content", in Woodfield.
- Noonan, H. (1984) "Fregean Thoughts", *The Philosophical Quarterly*.
- \_\_\_\_\_. (1986) "Russellian Thoughts and Methodological Solipsism, in Butterfield (1986b)
- Papineau, D. (1993) *Philosophical Naturalism* (Blackwell: Oxford)

- Perry, J. "Demonstratives", In *Demonstratives*, ed. Yourgrau.
- Platts, M. (1979) *Ways of Meaning* (London; RKP)
- Putnam, H. (1988). *Representation and Reality* (Cambridge Mass: MIT Press).
- \_\_\_\_\_. (1982) *Reason, Truth and History*.
- \_\_\_\_\_. (1975) *Mind, Language and Reality* (Cambridge: Cambridge University Press).
- Quine, W. (1956) "Quantifiers and Propositional Attitudes", *Journal of Philosophy*.
- Russell, B. "Philosophy of Logical Atomism", in *Logic and Knowledge*, ed., J. Marsh.
- \_\_\_\_\_. (1905) "On Denoting" *Mind*.
- Searle, J. (1983) *Intentionality* (Cambridge: Cambridge University Press).
- Segal, G. (1989) "Seeing What Isn't There" *Philosophical Review*.
- \_\_\_\_\_. (1989) "The Return of the Individual", *Mind*.
- \_\_\_\_\_. (1991) "In Defence of a Reasonable Individualism", *Mind*.
- Stalnaker, R. 1988. "What is in the head." In *The Nature of Mind*, ed., Rosenthal.
- Woodfield, A. (ed) *Thought and Object* (Oxford: Clarendon Press).