

THE PROBLEM OF COGNITIVE DYNAMICS*

by João Branquinho

Introduction

This paper is devoted to an examination of some aspects of the interesting topic of Cognitive Dynamics as it has been introduced by David Kaplan in his seminal essay ‘Demonstratives’ and further exploited in the sequel ‘Afterthoughts’.

I discuss two main sorts of approach to cognitive dynamics: the neo-Millian (or directly referential) approach, which I take as best represented in Kaplan’s views, and the neo-Fregean approach, which I take as best represented in Gareth Evans’s views. The upshot of my discussion is twofold. On the one hand, I argue that both Kaplan’s account and Evans’s account are on the whole defective (for different sorts of reason, of course); even though there are features of each of those views which seem to me to be along the right lines. On the other, and in spite of that, I claim that a broadly Fregean theory is still to be preferred since by positing semantically efficacious modes of presentation it is clearly better equipped to deal adequately with some important phenomena in the area. In particular, I argue that the notion of a memory-based demonstrative mode of presentation of an object (a spatio-temporal particular, a region in space, a period of time, etc.) turns out to be indispensable for the purpose of accounting for the persistence of an important range of mental states with propositional content over time.

The structure of this paper is as follows. In Section 1 one of the outstanding problems in cognitive dynamics is introduced and clarified. In Section 2 an immediate reaction to the problem, which I call the Natural Realignment Claim, is formulated and two different readings of it are isolated for further consideration. In Section 3 one of those readings, the Sufficiency Claim, is considered and rejected. Section 4 contains an examination of the remaining claim, the Necessity Claim, in connection with Kaplan’s and Evans’s views. Finally, in Section 5 those views are both regarded as inadequate and an alternative account is sketched, the Necessity Claim being eventually rejected.

1. The Problem

Cognitive Dynamics is in essence the study of attitude-retention and attitude-change. It is thus a discipline concerned with the conditions under which particular propositional attitudes - e.g. token beliefs, desires, and fears - can be said to persist over time, as well as with the conditions under which they can be said to have at some point ceased to exist. I assume for the purposes at hand a familiar, though not uncontroversial, general picture of propositional attitudes, namely the so-called relational view. On this view the attitudes are relational mental states, states which are at least binary in nature (so as to speak), holding between given subjects, on the one hand, and a certain breed of abstract objects usually referred to as *thoughts* or *propositions*, on the other; the latter are said to be the *contents* of the mental states and they are supposed to possess unrelativized truth-values, that is, to be semantically evaluated in an absolute manner. We shall be mostly concerned with attitudes that have *singular* thoughts, i.e. thoughts that are about *specific* objects, as their contents.

Let me then introduce what might be seen as one of the most challenging and delicate issues of cognitive dynamics. As formulated by Kaplan, the problem is mainly related to the phenomena of retention and change in those attitudes which have indexical thoughts as their contents, notably those thoughts which

might be expressed by utterances in given contexts of sentences containing at least one occurrence of an indexical referring expression. It is tacitly assumed that attitudes with allegedly eternal or non-indexical contents, e.g. presumably my fear that planes are likely to crash during take-offs or my belief that Barcelona is a very beautiful city, do not present any serious difficulties. Moreover, the species of indexicality whose investigation appears to be more interesting are especially those which are given in: (1) temporal indexical thoughts, in the sense of indexical thoughts about particular times; (2) spatial indexical thoughts, in the sense of indexical thoughts about particular places; and (3) perception-dependent indexical thoughts, in the sense of indexical thoughts about spatio-temporal particulars demonstratively identified by means of the several sensory modalities.

Cases where one is dealing with indexical contents of the above types are problematic and worth studying because they often involve what might be called a realignment in the linguistic means of expression of an indexical thought - on the part of a given thinker - as time goes by. In other words, there are situations in which the verbal expression of an indexical thought entertained by a thinker at a given time must, at a later time, be readjusted in a certain way by the thinker in order for the thought in question to be then entertained; so that one could presumably say that some particular attitude held at the earlier time towards the thought in question has been retained by the thinker at the later time, the very same thought being the object of the attitude on both occasions. Naturally, such readjustments are to be thought of as being operated in the linguistic means employed for the expression of the thoughts. In effect, it does not make much sense - at least in the light of the general picture of mental content we wish to endorse - to think of the thoughts as being themselves subjected to any sort of change or realignment. As we shall see, the problem is that certain linguistic means which *prima facie* suggest themselves as appropriate to the effect do not seem to capture, at least according to some views (including the account I will develop here), certain apparently intuitive facts about the cognitive structure of given thinkers.

Here is a general formulation of the problem, which for convenience I label *the central problem of cognitive dynamics*. Suppose that at a certain time t a given subject x holds a particular belief b with content **that p**, where the thought **that p** is a singular indexical thought. For example, suppose that, at t , x sincerely and assertively utters, or assents to, a token of a sentence-type S containing some indexical referring expression i and expressing, with respect to the context of use, the thought **that p**. We are then invited to consider the following questions. Under what circumstances should we say that, at a later time t' , x has retained her token belief b ? Notice that this question, taken as a question in *philosophical semantics*, is to be intended in the following sense. What sentence or sentences (if any) S' - presumably containing some indexical expression i' different from, but co-referential with, i - must x be disposed to assent to, or to assert, at t' so that we would be entitled to say that her particular belief b has been retained at t' ? Obviously, a natural and minimal requirement one should make here is that S' must have strictly the same content as S , i.e. it must have the thought **that p** as its content (relative to the context of use). Indeed, strict identity of mental content over time is clearly a necessary condition of attitude-retention.

And the related question concerning attitude-change is as follows. Under what sort of conditions should we say that, at t' , x has not retained her former token belief b and, for instance, has changed her mind with respect to such a belief (or to its content)? That is to say, what sentence or sentences (if any) S' must x be disposed to assent to, or to dissent from, at t' so that one could say that she no longer holds at that time her belief b ? As to this question, one should remark at once that there could be several ways by means of which one could say that, at t' , x has changed her mind about her original belief. For instance, she could at t' suspend judgement about the content of b , by neither believing the thought **that p** nor

believing the negation of that thought. Or she could at t' disbelieve the content of b , which would be tantamount to believing the negation of the thought that p (or so we might assume). Or she could find herself at t' in a position in which she simply does not believe that p at all, in the sense of failing to believe that p . In all these cases one would say that x has not retained at t' the particular belief b she held at t , one would say that her belief b has been dropped; for in all such cases she takes towards the thought that p at t' an attitude which is in some way antagonistic to the one she took towards the same thought at t .

The following sort of simple situation will illustrate the problem. Suppose that, at a certain time t , I sincerely and assertively utter, or assent to, a token of the sentence-type

(1) This room is hot now.

One would then say that at t I hold a particular belief about the temperature in the room where I happen to be at t ; the token belief might be formed on that occasion or it might already exist and just be linguistically manifested then. Now what sentence or sentences should I be prepared to accept at a later time t' , say three minutes later, in order to retain (or re-express) the particular belief I formed or voiced at t ? How should I readjust at t' the thought - relative to its verbal expression, of course - I entertained and believed at t ? Well, assuming that at t' I am still in the same room, a natural candidate would be a token of the sentence-type

(2) This room was hot three minutes ago;

or, allowing for a certain degree of vagueness, perhaps a token of

(3) This room was hot a moment ago.

To take another case, consider the notorious 'today'/'yesterday' example. Suppose that on a certain day, say d , Jones says

(4) Today is fine,

and believes it. What sentence or sentences should he be disposed to accept on the following day, $d+1$, so that one could say that he has then retained his previously manifested belief about the weather on d ? Again, a plausible choice would apparently be given in the sentence

(5) Yesterday was fine.

Analogous questions might be naturally raised with respect to sentences containing other categories of indexical terms, such as spatial indexicals (like 'here' and 'there') or perceptual demonstratives (like 'this pen' and 'that pen').

2. The Natural Realignment Claim

For convenience, let us call the above - *prima facie* intuitively sound - claim to the effect that sentences such as (2) (or (3)) and (5), taken with respect to the envisaged contexts of utterance, constitute choices which are appropriate for attitude-retention (or re-expression) the **Natural Realignment Claim**.

It is important to begin our reflection by noticing that such a claim might be given the following two

different readings. Take Jones's case, for instance. On the one hand, it might be held that Jones might be said to have retained on **d+1** the particular belief he held on **d** (by accepting then (4)) *only if* he would be disposed to accept (5) on **d+1**. That is to say, the claim is that a disposition to accept a token of the sentence in question on the later occasion is *necessary* for retention of the belief had on the earlier occasion. On the other hand, it might be held that Jones's disposition to accept (5) on **d+1** is *sufficient* for retention of the belief he held on **d** (by accepting then (4)). I shall refer to those distinct versions of the Natural Realignment Claim as the **Necessity Claim** and the **Unqualified Sufficiency Claim** (respectively).

It is instructive to spend some time arguing that the Unqualified Sufficiency Claim is false, that e.g. Jones's disposition to accept (5) on **d+1** is not sufficient for belief-retention. The reason is that from such an argument a stronger claim will emerge, the claim that the existence of such a disposition taken in conjunction with a certain set of additional suppositions is sufficient for belief-retention; call this the **Qualified Sufficiency Claim** (I shall deal with it later). Furthermore, our discussion of the Unqualified Claim will enable us to pick out a number of important features of attitude-retention.

3. The Sufficiency Claim

With a view to identifying the supplementary suppositions required by the Qualified Sufficiency Claim, let us then see why the unqualified version does not hold, e.g. why Jones's propensity to accept (5) on **d+1** is not - taken *per se* - sufficient for belief-retention. In what follows I mention three different ways by means of which such an insufficiency could be displayed.

First, suppose that on **d+1** - for instance as a result of a blow in the head - Jones forgets everything about the preceding day, especially what the weather was like then. Suppose further that on **d+1** - being aware of Jones's condition - a reliable friend gives him some information about **d**, including the information that it was a fine day (assume this was indeed the case); and suppose that Jones accepts this testimony without any reluctance. Then one might surely conclude that on **d+1** Jones would be disposed to accept a token of (5). Yet, under such circumstances, one would hardly say that he would have thereby retained on **d+1** the particular belief he held on **d** (by accepting then (4)). It seems that in order to secure belief-retention Jones's disposition to accept (5) on **d+1** must be causally connected in some way or other with his acceptance of (4) on **d**. Indeed, one would have to say that he is disposed to accept (5) on **d+1** *because* he accepted (4) on **d**; or that he believes today - i.e. **d+1** - that yesterday - i.e. **d** - was fine *because* he believed yesterday that yesterday was fine. And in the above sort of situation, such a causal connection - which, as the case suggests, may involve memory - is clearly absent.

Another way of establishing the same point, viz. the unsoundness of the Unqualified Sufficiency Claim, is this. Suppose that Jones accepts a token of sentence (4) at 11:58 p.m. on **d**; and that three minutes later, being unaware that midnight has already passed, he comes to accept a token of sentence (5). Of course, under such circumstances, with the word 'yesterday' Jones intends to refer to **d-1**, a day on which - he remembers (let us suppose) - the weather was also fine. Therefore, Jones accepts (5) on **d+1**; even though he would reject such a description of his situation if he were confronted with it. However, it is obvious that one could not claim that he has thereby re-expressed the token belief he held on **d**. Again, some kind of causal connection appears to be required for attitude-retention. And in those circumstances such a connection - which, as the case suggests, may involve the exercise of a tracking ability (the ability to keep track of a period of time as time passes) - is missing. In effect, in the envisioned situation one would not say that Jones accepts (5) on **d+1** *in virtue of* his previous acceptance of (4) on **d** (the reason

being that meanwhile he has lost track of day **d**).

One might be attracted by the general idea that *some* sort of internal continuity is constitutive of attitude-retention, the required causal connection being regarded as resting upon such continuity. And one way of fleshing out this idea, by specifying a particular form of internal continuity, would be roughly given in the following claim: **(a)**- necessarily, if a subject retains a given attitude with a given content for a given period of time then *throughout* the period in question she thinks that she is having the *same* attitude towards the *same* content. Or perhaps in the following, presumably weaker, claim: **(b)**- necessarily, if a subject retains a given attitude with a given content for a given period of time then *throughout* the period in question she does *not* think that she is *not* having the same attitude towards the same content. Clearly, the kind of continuity appealed to in these claims is not available in cases of the above kind. When Jones comes to accept **(5)** on **d+1** it certainly does *not* seem to him that he is believing then the same thing he believed on **d** when he accepted **(4)**: if asked he would claim to be believing a different thing. In particular, it certainly does *not* seem to him that he accepts **(5)** on **d+1** because he accepted **(4)** on **d** - for the simple reason that it does not seem to him that he accepts **(5)** on **d+1** at all.

However, I think that claims such as **(a)** and **(b)** pick out the *wrong* kind of internal continuity. Indeed, I am inclined to reject them on the grounds that they entail the thesis that attitude-retention is epistemically transparent to the subject; crudely formulated, this is the claim that a subject cannot be wrong about whether or not she retains a given attitude with a given content for a given period of time. Although I will not expand on the point here, I believe that the transparency thesis is mistaken and that, like most other mental phenomena, attitude-retention might well be epistemically opaque. It seems that it is perfectly possible for people to retain beliefs while thinking they have abandoned them, as well as to lose beliefs while thinking they have retained them.

Anyway, rejecting those claims does not entail rejecting the general idea about attitude-retention, that it should involve some sort of internal continuity. Moreover, other proposals for fleshing out the idea happen to be available, proposals which are arguably not committed to the transparency thesis. Let me mention two such proposals, both suggested by previous remarks and to which I shall return later. One of them brings out memory as one of the forms of interior continuity which are crucially present in attitude-retention. I take for granted that, in general, memory is not epistemically transparent to the subject. There clearly is a distinction between *remembering* and *seeming to remember*; for instance, it is possible to remember certain events while thinking that one does not remember them, or at least without thinking that one remembers them. The other proposal brings out the already noticed tracking capacity. One might remark with respect to Jones's case that he thinks that he is believing a different thing - when he accepts sentence **(5)** - in virtue of the fact that he thinks that he has correctly kept track of time, whereas in fact he has mistracked the days; wrong beliefs about content identity may thus be caused by wrong beliefs about tracking. Hence the continued exercise of an ability to keep track of an object, namely a day, as time flows may also provide us with a form of internal continuity of the kind needed for attitude-retention. Arguably, no commitment to the transparency thesis is forthcoming here either; for tracking might well be epistemically opaque. Indeed, a subject might be wrong about whether or not she has kept track of an object for a period of time; in particular, and this is perhaps more obvious with respect to perceptually presented items, she might have successfully tracked an object and yet believe she has lost track of it.

The notion of tracking an object has its clearest and primary application in the case of perception. Following Evans, one might maintain that an exercised ability to track a spatio-temporal particular over

time and/or space should be in general taken as being constitutive of that variety of attitude-retention one might call *P-retention*. Roughly, P-retention consists in a subject's re-expressing by means of perceptual demonstratives certain token attitudes previously held towards perception-dependent thoughts, for instance certain previous token perceptual beliefs (also expressed then with the help of perceptual demonstratives). I shall refer to the claim as principle **(K)** and give it the following formulation. Let the thought **that q** be a perception-dependent thought about some spatio-temporal item **v**. Then the following holds:

(K) Necessarily, if a subject *x* P-retains a token attitude *a*, of type *m* and with content that *q*, throughout an interval of time from t_0 to t_n , then *x* keeps track of *v* from t_0 to t_n .

Thus, suppose that on seeing a certain woman at a party I say 'She is ugly' and believe it. Then, the claim is this. In order to be able to P-retain on a later occasion during the party the particular belief about the woman I previously voiced by uttering that sentence - that is to say, in order to be able to re-express later that belief by means of a perceptual demonstrative (e.g. 'that woman') - I must somehow have kept track of the woman in question from one moment to the other (by means of some sensory modality, e.g. sight, or some combination of sensory modalities).

Now the foregoing remarks seem to point to an extension of principle **(K)** to cases of attitude-retention in which given subjects are able to re-express by means of temporal indexicals, e.g. 'yesterday' and 'a moment ago', certain token attitudes formerly held towards temporal indexical thoughts, for instance certain previous temporal beliefs involving 'today'-thoughts or 'now'-thoughts. Let us refer to such a variety of attitude-retention as *T-retention*. Then one might formulate as follows the envisaged temporal analogue of principle **(K)**. Let the thought **that r** be a temporal indexical thought about a particular time (instant or interval) **u**. Then the following holds:

(K)* Necessarily, if a subject *x* T-retains a token attitude *a*, of type *m* and with content that *r*, throughout an interval of time from t_0 to t_n , then *x* keeps track of *u* from t_0 to t_n .

Thus, for example, in order for Jones to be able to T-retain on **d+1** - or to re-express on **d+1** by means of a temporal indexical - the particular belief he held on **d** about **d** (by accepting then a token of **(4)**), he must keep track of **d** from **d** to **d+1**. And an analogue of principle **(K)** could also be formulated to cover cases of what one might call *S-retention*, cases in which a subject re-expresses by means of spatial indexicals, e.g. 'over there', certain token attitudes formerly taken towards spatial indexical thoughts, e.g. certain previous spatial beliefs involving 'here'-thoughts.

Principles **(K)** and **(K)*** will be examined later (Section 5). Let me now go back to the Unqualified Sufficiency Claim and introduce a third way of displaying its inadequacy. One might imagine a situation in which our subject Jones simply changes his mind on a certain occasion between his acceptance of **(4)** on **d** and his acceptance of **(5)** on **d+1**. Suppose that Jones accepts **(4)** at some time **t** during the afternoon of **d**, and that in the evening of **d** the weather gets temporarily pretty miserable so that at **t'** he comes to doubt that the weather is fine on **d**. Then at some time **t''** on **d+1**, after having given some reflection to the topic of the weather on the preceding day, he ends up accepting **(5)**. One may assume that at **t''** Jones remembers what the weather was like on **d** and that he has successfully kept track of the days from **t** to **t''**. However, I guess that in such circumstances one would not be entitled to say that Jones has retained at **t''** the belief he formed at **t**, or that this belief persists throughout the period that goes from **t** to **t''**. Maybe there could be a sense in which one would say that such a belief is *resumed* at

t''. Yet, if token beliefs are viewed (as they should be) as particulars, and not as universals, then they are as such non-repeatable items and so the idea is utterly meaningless that they might cease to exist at one time and yet reappear at another. Therefore, at best, to say that Jones's belief is resumed at **t''** would only make sense if 'belief' were taken as meaning *type* of belief (in some appropriate sense of 'type') and not token belief; so what one should say in the end is that Jones's token belief about the weather on **d** has not been retained at **t''**.

The following general conclusion might be extracted from our reflection on the 'today'/'yesterday' case; and parallel results would surely hold for other pairs of temporal indexicals such as e.g. 'now'/'three minutes ago', for pairs of spatial indexicals such as e.g. 'here'/'there', and for pairs of perceptual demonstratives such as e.g. 'this pen'/'that pen'. Given an acceptance of a token of **(4)** on a day **d**, a subject's disposition to accept a token of **(5)** on **d+1** would not by itself guarantee the persistence of the belief held on **d**. The reason is that at least one of the following conditions might not be satisfied: **(i)**- the subject remembers the weather on **d**; **(ii)**- he keeps track of the days from **d** to **d+1**; **(iii)**- he does not change his mind about his earlier belief. Therefore, what I take to be the Qualified Sufficiency Claim might now be given the following formulation: the presence of that disposition, in conjunction with such conditions as **(i)**, **(ii)** and **(iii)**, would be sufficient for attitude-retention.

4. The Necessity Claim

The questions I would like to address in the remainder of the present essay are these. Is the Natural Realignment Claim - construed as the Qualified Sufficiency Claim - a sound claim? And is the Natural Realignment Claim - construed as the Necessity Claim - a sound claim? Let us begin by considering the latter issue.

Recall that the Necessity Claim is the claim that, given an acceptance of a token of sentence **(4)** on **d**, a propensity to accept a token of sentence **(5)** on **d+1** would be necessary to retain then the belief held on **d**. Now it might be informative to be told that this claim *might* be subscribed to by an advocate of a directly referential account of indexical content. That is to say, the conceptual machinery usually available on such an account might be employed so as to accommodate the Necessity Claim in a natural way. Indeed, a direct reference theorist might come to endorse the idea that *only* by being disposed to accept a sentence such as **(5)** on **d+1** would our subject be able to retain the particular 'today'-belief he held on **d**. For she might insist that, at least in the case of singular indexical belief, belief-retention should be regarded as being sensitive, not only to believed **contents**, but also to **ways of believing contents -characters**, in Kaplan's terminology, or **roles**, in Perry's terminology. Thus, in order for Jones to re-express his 'today'-belief on the following day it would not be enough for him to accept then the same content, even if all the conditions mentioned in the previous Section were met. Note that this content would be (roughly speaking) the neo-Russellian proposition consisting of the day **d** itself and the attribute of Fineness. Yet, Jones could believe such a proposition on **d+1** under a variety of ways, including non-indexical ones such as those involving date indicators like 'July 3, 1997' or 'John's birthday'. And it might be reasonably argued that not every way of believing a neo-Russellian proposition is appropriate to secure belief-retention. Accordingly, given that on **d** the proposition in question is believed by the subject under the character (of) 'today' (*inter alia*), the direct reference theorist might impose the following condition on belief-retention. She might demand that that character, i.e. the "old" singular character, be readjusted in such a way that: (1)- the "new" singular character determines the same neo-Russellian content, viz. the same object; *and* (2)- the new character is suitably related to the old character, in the sense of belonging to what might be loosely called the same **range** of

singular characters or ways of believing. 'today', 'yesterday', 'tomorrow', 'two days ago', etc., would form a single range of singular characters in that sense, and 'here', 'there', 'over there', etc., would belong in another range. The Necessity Claim might then be naturally vindicated: only by entertaining on **d+1** the 'yesterday'-way of believing the above neo-Russellian proposition would Jones be able to retain his previous 'today'-belief.

However, the criticism might be levelled against such a directly referential account that it would fall short of providing us with any kind of *interior* continuity. On the one hand, neo-Russellian propositions are by definition something that is, in general, wholly external to their thinkers. On the other, it seems unlikely that the notion of a character, or a way of believing a content, might be adequately used to the effect. On **d+1** Jones is required to accept the same content under a suitably related character (in the above sense). However, it is still unclear what relation should obtain between the old and the new character so that it could yield the interior continuity involved in attitude-retention. Hence, lacking the appropriate kind of conceptual apparatus, at least to the extent that the semantic level of explanation is concerned, the direct reference theorist seems to be left with the problem of accounting in some way or other for such continuity. Thus, if we are right, she seems to be left with the general problem of accounting for attitude-retention. Nevertheless, I think this dismissal would be too quick. Indeed, one possible move available to the direct reference theorist would be to invoke at this point the already mentioned tracking capacity and take it as providing us with the sort of connection among characters which is involved in belief-retention. The idea is that underlying a given range of indexical ways of believing of the envisaged types (temporal, spatial, perceptual) is the exercise of an ability to keep track of an object over space and/or time. This would give us the needed "internal" element, while being perfectly consistent with the coarse-grained contents posited by the account: structures of objects and properties or relations. Of course, one might at this point protest that, by introducing ways of believing (modes of presentation, senses) which are far from being semantically inert, and notwithstanding the adoption of a purely referential notion of indexical content, the emerging theory would turn out to be more Fregean than Millian. But that is a different issue, irrelevant to our present concerns.

On the other side, the "interior continuity" issue would not represent a problem for neo-Fregean approaches to indexical content, especially the particular account advanced by Evans. The reason is that the considerations usually employed on such an account to establish the result that Jones's acceptance of **(5)** on **d+1** would, under normal circumstances, manifest the persistence of the particular content believed by him on **d** (when he accepts **(4)**), are such that they provide us at once with the desired interior continuity. In effect, on that brand of Fregeanism, it is claimed that the particular thought Jones believes by accepting **(4)** on **d** is strictly identical to the one he believes by being disposed to accept **(5)** on **d+1**. Because both thoughts are composed out of the very same particular way of thinking of an object (a day); and such a way of thinking is taken to be a temporally extended mode of presentation of the day **d**, one which consists in a way of keeping track of **d** from **d** to **d+1** which Jones employs in thought. Thus, a particular form of internal continuity, the one noticed before and given in the exercise of an ability to keep track of a period of time as it recedes into the past, is in this manner incorporated into indexical content itself. And the neo-Fregean theorist would be able to conclude that a subject's disposition to accept **(5)** on **d+1** is necessary for her to retain the content believed on **d**. For, on his view, if the subject keeps track of the day **d** from **d** to **d+1** then she is bound to have such a disposition on **d+1**, and if she wants to retain then the content of her earlier belief, then she must exercise such an ability. An argument along the following lines could then be mounted with a view to establishing the Necessity Claim. If the above result holds, i.e. if a disposition to accept **(5)** on **d+1** is necessary for preservation of content, then

such a disposition is necessary for belief-retention; since, in general, belief-retention requires identity of content believed. Alternatively, a different argument towards the same conclusion could be set out in the following way. One might invoke the principles introduced earlier on about tracking objects and claim that, in general, retention of demonstrative or indexical beliefs about particular objects requires that the subject keep track of the objects thought about throughout given periods of time. Thus, in particular, Jones retains on **d+1** his belief held on **d** about **d** only if he keeps track of **d** from **d** to **d+1**. On the other hand, as claimed above, in order to be able to keep track of **d** from **d** to **d+1** he must be disposed to accept **(5)** on **d+1**. Therefore, the conclusion is forthcoming that such a disposition would be necessary for belief-retention.

Summing up the preceding discussion, we could say that the Necessity Claim might be accepted **both** by a neo-Fregean theorist and by a direct reference theorist in so far as they might both endorse the following general line of argument (though on different grounds). *Premise 1*: Tracking an object over time and/or space is necessary for retaining on later occasions singular indexical beliefs expressed on earlier occasions by means of sentences of certain kinds ('today'-sentences, 'here'-sentences, etc.). *Premise 2*: A disposition to accept sentences of certain kinds ('yesterday'-sentences, 'there'-sentences, etc.) on the later occasions is necessarily involved in tracking. *Conclusion*: Therefore, such a disposition is necessary for belief-retention.

Now the strongest objection I know to the Natural Realignment Claim - or at least to its construal as the Necessity Claim - is one that comes from the "direct reference" side: Kaplan's criticism in 'Demonstratives'. Kaplan dismisses as unsatisfactory such an answer to the central problem of cognitive dynamics; and, as far as I can see, he does not provide us with any alternative solution to the problem, leaving it open. His argument might be represented as the following sort of *reductio*. If, in order to be able to retain on **d+1** his former belief, Jones must replace 'today' with 'yesterday' and accept **(5)**, then if he were someone like Rip van Winkle - someone who systematically loses track of time and never knows his own position in time - it would be utterly impossible for him to retain that belief; for the simple reason that he could never be in a position to accept, or to be disposed to accept, a sentence such as **(5)** (in a sincere and reflective manner, of course). Naturally, Kaplan takes this as a rather implausible consequence of the claim under discussion. Thus, what Kaplan seems to reject is some such idea as the idea that in order to retain demonstrative or indexical beliefs about a given object for a period of time a subject must keep track of the object in question throughout that period of time; in other words, Kaplan seems to reject some such principles about retention as **(K)** and **(K)***, and in general Premise 1 of the above argument for the Necessity Claim.

The strongest rejoinder I know to Kaplan's argument is Evans's in his influential essay 'Understanding Demonstratives'. Evans rejects Kaplan's verdict and takes the above sort of consequence as being perfectly acceptable. He argues that an ability to keep track of time - which, as we have seen, manifests itself in a propensity to accept **(5)** on **d+1** given an acceptance of **(4)** on **d** - necessarily underlies Jones's retention of his original belief. If Jones were like Rip van Winkle and that ability were missing, there would be no belief-retention. Rather, the subject would be under the illusion of having retained a certain belief.

How should one assess such a dispute? First, notice that the case presented by Kaplan is an extreme one. And I am inclined to think that Evans is probably right when he holds that a subject who *systematically and massively* mistracks time - e.g. by sleeping for too many consecutive days - could hardly be in a position to retain temporal beliefs; at least, he could hardly retain temporal beliefs involving 'today'-

thoughts, ‘now’-thoughts, and the like. In effect, maybe the right thing to say about someone like Rip van Winkle is that such a person would be unable to *form* any such temporal beliefs in the first place - and hence to retain them, of course. However, there is no need for such an extreme case. Indeed, an analogous point could be made with respect to cases in which we deal with subjects who only locally or temporarily lose track of time. Thus, let us go back to Jones’s story. Suppose that he accepts (4) at 11:58 p.m. on **d** and that, three minutes later, he is unsure whether midnight has passed; suppose further that he has no means to check the time. Hence, Jones would not be in a position to be disposed to accept (5) at 00:01 a.m. on **d+1**, such a disposition being grounded on an ability to keep track of a day which he happens not to be able to exercise in the situation described. Therefore, according to the Necessity Claim, it seems that Jones would be prevented from retaining at 00:01 a.m. on **d+1** the belief he held at 11:58 p.m. on **d** (even assuming that he does not misremember the weather conditions on **d**). Yet, I think that there apparently is - at least at a pre-theoretical level - a sense in which at 00:01 a.m. on **d+1** Jones still continues to believe what he believed three minutes earlier. That is to say, it seems to me that there is an intuition to the effect that it should be possible for Jones to retain then the belief held on the day before about the weather on that day; even if he is *ex hypothesi* not able to think of the day in question as **yesterday**, or as **the day which immediately precedes today**, or to identify it by means of knowledge of his own position in time. (Naturally, parallel considerations could be made with respect to cases of retention of spatial or perceptual beliefs.)

5. Memory and Belief-Retention

The sort of approach to the central problem of cognitive dynamics I am willing to endorse, and whose main ideas are subsequently sketched, is one which is both in partial agreement with Evans's views and in partial agreement with Kaplan's views.

On the one hand, *pace* Evans, I have reasons to believe that the Necessity Claim is not in general sound. In particular, as pointed out, I would agree with Kaplan's verdict that in the above sort of circumstance - in spite of not being disposed to accept (5) on **d+1**, and hence in spite of not being in a position to keep track of **d** in a certain canonical way, viz. as **the preceding day** - Jones might still be seen as having retained his earlier belief. What sort of retention would that be, and how might Jones manifest it, are questions I take up below. On the other hand, I have reasons to believe that claims such as principles (K) and (K)* still impose plausible constraints upon a subject's ability to retain and re-express in a certain way propositional attitudes with demonstrative or indexical contents. Thus, *pace* Kaplan, I think that the notion of keeping track of an object over time and/or space - in the perceptual case, the notion of following or monitoring a public item as it moves around, we move around, and time passes - should still be seen as playing an important role in an adequate account of attitude-retention. Our rejection of the Necessity Claim should not be regarded as being inconsistent with our acceptance of such principles as (K) or (K)*. Indeed, I think that there is more than a kernel of truth in Evans's doctrine that in order to retain or re-express *in a certain manner* a previous demonstrative or indexical belief a subject must keep track of the object thought about; although I have some reservations concerning his doctrine that ways of tracking objects - for instance, the one underlying an appropriate succession of beliefs involving ‘today’, ‘tomorrow’, and ‘yesterday’ - are themselves to be taken as component parts of demonstrative or indexical contents.

Nevertheless, I maintain that the Necessity Claim does not hold. As claimed before, the reason is that I accept the idea, which I see as being grounded to a large extent on sound intuitions, that it is possible for a subject to retain, and also to re-express, on **d+1** a belief held on **d** (by accepting (4)) without having

then any disposition to accept (5). One might perhaps represent the basic mistake made by a proponent of the Necessity Claim as consisting in a tacit endorsement of some strengthened and unrestricted forms of principles such as (K) and (K)*. According to the stronger principles in question, a subject's exercise of an ability to track a given object over time and/or space would be required for the subject to retain *tout court* perceptual or temporal beliefs previously held about the object in question. Whereas, according to the formulations proposed and which I take as plausible, the weaker principles (K) and (K)* are restricted respectively to cases of P-retention, i.e. re-expression of former perceptual beliefs by means of perceptual demonstratives, and to cases of T-retention, i.e. re-expression of former temporal beliefs by means of temporal indexicals. In a different terminology, one could say that those principles operate by limiting retention and re-expression to ways of believing available *within* a given range of ways of believing. So one might perhaps view the Necessity Claim as resting upon an assimilation of attitude-retention in general to T-retention, in the temporal case, and to P-retention, in the perceptual case.

Sed Contra such an assimilation, I am prepared to hold the view that not every instance of retention *tout court* is to be taken, for example, as an instance of T-retention (even when the objects of retention are temporal beliefs). Thus, if Jones mistracked time in the way described before, then - in the light of principle (K)* - what he would not be in a position to do at 00:01 a.m. on $d+1$ is to T-retain the particular belief he held at 11:58 p.m. on $d+1$ when he accepted (4). In other words, he would not be able to re-express then such a belief by using a temporal indexical such as 'yesterday'. Nevertheless, it does not follow that it would be impossible for him to retain, or even re-express, *tout court* that particular belief. (Analogous observations would hold for P-retention and S-retention.)

The account we have developed so far already provides us with the sort of result we need to make sense of such a possibility. Indeed, it is almost a truism that there is an intimate connection between attitude-retention and memory. For simplicity, let us confine our attention to belief-retention. Then the following principle captures a fundamental aspect of the intended connection by stating that propositional memory entails belief-retention:

(M) Necessarily, a subject x retains a token belief b , with content that p , throughout an interval of time from t_0 to t_n provided that at any time t_i in the interval x remembers that p .

I take it that (M) is in itself a rather plausible principle, standing in no need of being established by any substantive argument. Of course, there are cases in which one would say such things as "She perfectly remembers that so and so, but does not believe it". But those cases are not genuine counter-examples to the principle as they are likely to rest upon irrelevant notions of belief, such as the one on which one is said to fail to believe that p when one wishes it were not the case that p . Alternatively, for those who find plausible the claim that knowledge entails belief (the claim that, necessarily, if x knows at t that p then x believes at t that p), principle (M) might be seen as a logical consequence of that claim taken in conjunction with the hardly disputable claim that propositional memory is a species of propositional knowledge (the claim that, necessarily, if x remembers at t that p then x knows at t that p).

Principle (M) is a first-order principle, in the sense that it does not involve any iteration of propositional attitudes, any attitudes about attitudes. What it requires the subject to remember is precisely that which she believed - the content of her memory is nothing over and above the content of her belief. In particular, the principle does not require the subject to have a memory about a belief, to remember that she *had* such a belief (or even to remember *having* such a belief). Therefore, principle (M) should not be conflated with the following sort of claim also connecting memory and belief-retention:

(M)* Necessarily, if a subject x retains a token belief b , with content that p , throughout an interval of time from t_0 to t_n , then at any time t_i in the interval x remembers that x believes that p .

(M)* is a higher-order principle, as it requires the subject to remember **that she believed** a certain thing. Anyway, I think that principle **(M)*** is false even if one restricts its application to creatures which possess the concept of *belief*. I take it that it is perfectly possible for a (rational) subject who at a given time still holds a certain belief not to be aware then that she continues to have the belief; for example, she might mistakenly think that she is then believing a different thing. It follows on reasonable assumptions that the subject could not be in a position to remember that she still has the belief in question. Hence, I do not think that retention of a previous belief is a matter of remembering that one still has the belief in question (the same would hold for the other attitudes).

Assuming that principle **(M)** is along the right lines, one could handle cases such as Jones's as follows. Supposing that on **d+1** he remembers what the weather was like on **d**, in the sense of **remembering that** the weather on **d** was such and such, one might say that he would be then in a position to retain the belief had on **d**, such a retention being secured by memory. Accordingly, what one should take as being the extreme case here is not the case of a massive loss of a capacity to keep track of time (Rip van Winkle's situation), but rather that of a massive loss of memory, i.e. the case of amnesia. Hence, what one ought to say is that temporal beliefs such as the ones we have been discussing could hardly be retained by an amnesiac (*mutatis mutandis* for perceptual and spatial beliefs of the envisaged kinds).

To sum up our discussion of the Necessity Claim, if the foregoing remarks are correct then it is reasonable to assert the following. Although Jones is *ex hypothesi* unable on **d+1** to keep track of **d**, i.e. to think of **d** as **yesterday**, he still might be said to have retained on **d+1** his former belief about **d** in a certain way, namely by means of memory, and not in virtue of the particular position he occupies in time or of his knowledge of such a position. I conclude then that there are solid grounds on which one might regard as unacceptable the claim according to which a disposition to accept **(5)** on **d+1** is necessary for a thinker to retain or re-express then a belief she had on **d** by accepting **(4)**. In the light of what was urged before, the existence of such a disposition is not necessary for belief-retention; and it is not sufficient either (for different but rather obvious reasons).

Assuming thus that what I have called the Natural Realignment does not constitute a necessary condition for belief-retention, one is still left with the problem of specifying what sort of linguistic readjustment (if any) could Jones make on **d+1** with a view to expressing then the persistence of the particular belief he had on **d**. A tentative proposal, suggested by some of the preceding observations, might be sketched in the following way. When Jones accepts **(4)** on **d**, he might as well accept a token of a sentence-type such as

(6) This day is fine;

where the demonstrative phrase 'this day' is associated in the context with a perceptual demonstrative mode of presentation of the day thought about, viz. **d**. In effect, it seems plausible to regard 'today' - taken in its pronominal use - as having a genuinely demonstrative sense, so that an utterance of **(4)** on a certain day might be seen as expressing the very same thought as an utterance of **(6)** on the same day. Now if such a suggestion is plausible, then a way by means of which a thinker like Jones could re-express on **d+1** the belief held on **d** (by accepting **(4)** then) would be to accept, or to have a disposition to accept, on **d+1** a token of a sentence-type such as

(7) That day was fine.

The demonstrative phrase 'that day' would be here taken as expressing in the context a memory-based demonstrative mode of presentation of **d**, i.e. a way of thinking of a certain day anchored upon a memory demonstration of the day in question. Reference to a particular day would be in this way secured by memory; and one might still say that the thinker knows which day is in question, or that he is in a certain sense acquainted with the object thought about, since he thinks of that day as the day on which certain remembered events took place. The idea is that belief-retention would be secured by means of the following sort of realignment of mode of presentation: the primitive perceptual demonstrative mode, the one associated with 'this day' (or 'today'), would be readjusted into a memory-based demonstrative mode, the one associated with 'that day', the same object (day) being presented by both modes.

If one were a strict Fregean and regarded modes of presentation as being locatable at the level of semantic or mental content, then one would have to argue to the effect that such a readjustment of mode of presentation does not constitute a literal change and does not create a *new* mode of presentation. Accordingly, from that point of view there would be a single temporally extended demonstrative mode and a hence a single Fregean thought would be expressed both by **(6)** and **(7)** (taken in the intended contexts). But that is far from being a compulsory move. Moreover, I think it is extremely difficult to make sense of the claim that a single mode of presentation is involved in such cases - there just seem to be two distinct ways of thinking about an object; and that claim is absolutely needed to guarantee identity of content. So perhaps one should prefer to be a less strict Fregean and favour a view on which modes of presentation, in spite of being external to semantic or mental content, are nevertheless semantically efficacious. Roughly speaking and for present purposes, modes of presentation (or ways of believing contents) are semantically efficacious in the sense that ascriptions of attitude-retention (and of attitude-change), in particular the correctness or incorrectness of such ascriptions, are sensitive to the identity and existence of modes. Moreover, on such a Fregean view it would be rather natural to take semantic or mental content as being purely referential or coarse-grained (in the neo-Russellian sense of complexes of objects and properties or relations). The above kind of readjustment of mode of presentation might then be regarded as constituting a literal change, in the sense that two demonstrative modes would be involved there: the "old" perceptual mode and the "new" memory-based mode. Yet, on the view in question one would still have a single, purely referential, thought; and strict identity of content could be accounted for in a presumably less troublesome way. Thus, belief-retention would be secured, not at the level of content as on the strictly Fregean view, but at the level of semantically efficacious modes of presentation: only certain modes, modes endowed with certain features and related in a certain way, would be appropriate to the effect.

There is still one issue I would like to briefly address. So far, we have restricted our attention to cases where subjects are able to retain or re-express true beliefs, i.e. beliefs with true contents. However, what should one say about retention of false beliefs? Clearly, these beliefs cannot be retained in the way suggested in Principle **(M)**, i.e. on the basis of propositional memory of their contents. The problem is not that the case of false belief invalidates the principle; it obviously does not. The problem is that the principle is not doing any explanatory work, it does not tell us how retention of false beliefs is achieved. On the other hand, for reasons already mentioned, I do not think that an appeal to higher-order principles such as **(M)*** would be a good idea. I believe that the above proposal, notably *via* the notion employed there of a memory-based demonstrative mode of presentation of an object, already contains the resources which one needs to accommodate the case of retained false belief. Propositional memory will not do. But

what might be loosely called "objectual" memory, or memory of things, will do. What I have in mind here are especially memories of items which are no longer perceptually accessible - some of them, e.g. past days, are even inaccessible in principle - but with which one had previous perceptual encounters. Take Jones's case again and suppose everything is as before except that the weather on day **d** is nasty. How could he retain on **d+1** the particular 'today'-belief he held on **d** by accepting (4) then? I think he would be able to do it on the basis of his thinking that a certain *remembered day* - in the sense of a day demonstrated by means of a memory - was fine. In other words, he would be able to do it on the basis of entertaining then a memory-based demonstrative mode of presentation of the day **d** and (mistakenly) thinking that **d** - thus presented - was fine. Therefore, he might naturally re-express that belief in the way suggested before by means of an acceptance of a sentence such as (6) (construed as indicated).

Finally, let us turn our attention to that reading of the Natural Realignment Claim, which we have labelled the Qualified Sufficiency Claim. Recall that this claim is to the effect that if a subject who accepts (4) on **d** is disposed to accept (5) on **d+1** - by keeping track of **d** in the "canonical" way - then she retains on **d+1** the belief held on **d**, provided that the following conditions are also satisfied: (i)- she remembers that the weather was fine on **d**; and (ii)- she does not change her mind about her former belief. What should one say about such a claim? Well, I take it as a perfectly acceptable one. In effect, at least with respect to cases of belief-retention, it turns out to be a simple logical consequence of our principle (M) (by means of the rule of addition of premises). So there seems to be no reason to reject the view that, by being disposed to accept (5) on **d+1**, the subject does not form a new belief, different from the one held on **d** by accepting (4) (assuming that the above conditions are fulfilled). The basic idea we have tried to introduce is just that such a belief might be retained and re-expressed by the subject on **d+1** in a different way.

João Branquinho

Department of Philosophy

Faculdade de Letras de Lisboa

Cidade Universitária

1699 Lisboa Codex – Portugal

References

Evans, Gareth

- **1981:** 'Understanding Demonstratives' in H. Parret and J. Bouveresse (eds.), *Meaning and Understanding*, Berlin, W. de Gruyter. Reprinted in G. Evans, *Collected Papers*, Oxford, Clarendon Press, 1985, 291-321.

Frege, Gottlob

- **1967:** 'The Thought: A Logical Inquiry', translated by Anthony and Marcelle Quinton, reprinted in P. F. Strawson (ed.), *Philosophical Logic*, Oxford, Oxford University Press, 17-38.

Hoerl, Christopher

- **1997:** 'Cognitive Dynamics: An Attempt at Changing Your Mind' in J. Dokic (ed.), *Cognitive Dynamics. European Review of Philosophy, Vol. 2*, Stanford, CSLI Publications, 141-158.

Kaplan, David

- **1988a:** 'Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals' in J. Almog, J. Perry and H. Wettstein (eds.), *Themes from Kaplan*, Oxford, Oxford University Press, 481-563.

1988b: 'Afterthoughts' in J. Almog, J. Perry and H. Wettstein (eds.), *Themes from Kaplan*, Oxford, Oxford University Press, 565-614.

Peacocke, Christopher

- **1983:** *Sense and Content*, Oxford, Oxford University Press.

Perry, John

- **1980:** 'A Problem about Continued Belief', *Pacific Philosophical Quarterly* **61**, 317-332.

1997: 'Rip van Winkle and Other Characters' in J. Dokic (ed.) *Cognitive Dynamics. European Review of Philosophy, Vol. 2*, Stanford, CSLI Publications, 13-40.

Richard, Mark

- **1990:** *Propositional Attitudes. An Essay on Thoughts and How We Ascribe them*, Cambridge, Cambridge University Press.

Williamson, Timothy

- **1995:** 'Is Knowing a State of Mind', *Mind* **104**, 533-565.

ã 1998 João Branquinho