



CHRISTOPH HOERL

Cognitive Dynamics: An Attempt at Changing Your Mind

In “Demonstratives” David Kaplan introduces the problem of cognitive dynamics, the question of what it means to have retained a belief over time which one would have expressed on a previous occasion by using an indexical expression. In a footnote, Kaplan points out that an investigation of the phenomenon of cognitive dynamics might start with the following question:

[W]hat does it mean to say of an individual who at one time sincerely asserted a sentence containing indexicals that at some later time he has (or has not) *changed his mind* with respect to that assertion? (Kaplan 1989: 538n).

In this investigation, I wish to take up Kaplan’s suggestion. I shall choose the phenomenon of changing one’s mind as my starting point and work my way from an analysis of this phenomenon to an understanding of cognitive dynamics. Of particular importance for this investigation will be the precise structure of the temporal reasoning a subject has to engage in if she wants to change her mind about something. I shall defend a theory according to which the following three conditions must be satisfied in order for it to be the case that you have changed your mind about x :

- (1) First you believed of x that it was φ , now you believe of x that it is not- φ .
- (2) You stopped believing of x that it was φ before you started believing of x that it was not- φ .
- (3) You know that (1) and (2).

My task will be to show why I think (1)–(3) are necessary within an analysis of the phenomenon of changing one's mind.¹ (3) entails (1) and (2), but not vice versa, and the main part of this investigation will be taken up by a discussion of what (3) adds to (1) and (2). My basic argument might be described as follows: As they figure in both (1) and (2), your beliefs concerning *x* are *referentially transparent*, i.e. you need not be aware of the fact that they concern the same object. The role of (3) lies in making sure that it is in fact *epistemically transparent* to you that both beliefs are concerned with the same object.² I shall argue that it is in this condition that the analysis of a change of mind connects with a theory of cognitive dynamics, a theory according to which the sameness of one's beliefs over time can be an epistemic given.

1.

There are many alterations in 'cognitive state' which don't qualify as changes of mind.³ (1) serves to rule out three types of such alterations. Under the first of these types fall alterations which are due to the appearing or disappearing of beliefs. No change of mind occurs when somebody acquires a belief about a thing she didn't even know about before. Similarly, forgetting is not a change of mind. (One solution (1) leaves open, however, is that a change of mind might be a co-occurrence of forgetting one belief and acquiring a different one.) A change of mind is a transition from one belief state to another, not merely a transition from a belief state to some other state or vice versa.

A second group of alterations ruled out by (1) concerns those cases in which the subject's beliefs held at different times concern different referents. Of course, a subject can come to possess different beliefs about the weather, for instance, just in virtue of looking out of the window at different times. But this has nothing to do with a change of mind. As Baier puts it: "Varying one's intentional activity to fit varying circumstances, like updating one's

1. I hesitate to call them sufficient conditions, since I think that a satisfactory analysis of the phenomenon of a change of mind should be pitched at the level of the justification rather than that of the dynamics of the change. Nevertheless, I assume that such an analysis would entail conditions (1)–(3).
2. I adopt the distinction between referential and epistemic transparency from Boghossian (1994) who draws upon Quine (1960) in choosing this terminology.
3. See Baier (1978) and Dennett (1978) for some examples.

version of the current circumstances, requires no change of mind" (Baier 1978: 161). As these considerations make clear, however, we shall have to restrict the domain of objects '*x*' can stand for in a way that isn't yet explicit in (1).⁴ Basically, we have to make sure that the object the beliefs are about will only be considered in so far as it doesn't undergo change with respect to the property salient for the truth of the subject's beliefs. Of some things this will be true by definition, such as particular times, or time-slices of material objects. However, we can also allow things into the domain under consideration which don't change as a matter of contingent fact, such as material objects in a particular period of their lifetime. For the sake of this discussion, all we need to do as far as beliefs about enduring things are concerned is to count changes *in* the referent as changes *of* the referent.

There is a third group of cases ruled out by (1), albeit one which is not as clearly defined as the other two. It concerns those alterations in one's cognitive state in which the subject does not move from a belief containing one predicate to a belief containing a contradicting predicate. According to recent externalist or anti-individualist theories,⁵ an alteration in one's cognitive state can be due to the fact that one has moved from one linguistic community to another. If the two communities have different linguistic norms, the thought which one used to express by uttering a sentence in the first community (and as a member of that community) can differ in meaning from the thought one expresses in the second community (and as a member of that community) by uttering the same sentence. This variation in meaning can happen without one being aware that this is the case. A similar phenomenon can be found in more common cases. Think of someone who used to think of one of his friends 'He is lazy' but has since become more tolerant and now thinks of the same friend 'He is lacking energy'. Has he changed his mind? I would say that the answer, in both cases, turns on the question of whether the subject conceives of the predicates as contradicting each other or not. In the case of the subject switching from one linguistic community to the other, it is precisely part of the externalist's story

4. Note that this does not mean giving up the *de re* specification of the beliefs in question, because it is precisely a restriction on the *res* they can be about.
5. Cf. esp. Burge (1988). For a discussion of this phenomenon cf. Boghossian (1994). It will become clear that I disagree with Boghossian on the implications this has for self-knowledge.

that in order to be in a mental state with a certain (externally individuated) content the subject does not need to be able to distinguish the content of this mental state from that of states she could be in were the circumstances different. Similarly, the person adopting a more tolerant view of his friend might do so against the background of a new outlook on people's behaviour in general (mellowing with age), a revision of her previous views which does not necessarily involve explicit rejection of them.

Whether there is a change of mind in these cases turns on the question of what the thinker's holding the new belief implies. As an analogue, we may consider a similar question arising with respect to the implications of what people say in a conversation. It will rarely be the case that a speaker will straightforwardly contradict a statement another speaker has made. Rather, we will often find situations in which upon one speaker's uttering a sentence like, for instance, 'Bill is lazy' another speaker will reply with a statement such as 'Bill is lacking energy'. What the second speaker says does not by itself determine whether she disagrees with the first speaker's assessment of Bill or not. She can indicate disagreement by prefixing her statement with the word 'No', or by uttering it in a particular tone of voice, thereby making clear that she intends her utterance to carry the implication of rejecting the first speaker's judgement. But it is by no means clear that the utterance necessarily carries such an implication, and it can sometimes be part of the goal of a conversation for the speakers to reach agreement on what they agree or disagree on. Returning to the case in which someone entertains different beliefs at different times, we can therefore say that her holding those beliefs only constitutes a change of mind if she conceives of the predicates which figure in them as mutually incompatible. The thinker's holding those beliefs however, will not always commit her to making such a judgement, moreover, she might not even be able to do so.

2.

I have said that (1) draws a distinction between changes of mind and certain other alterations in a thinker's cognitive state. However, strictly speaking, (1) does not ensure that any alteration in the thinker's cognitive state has occurred at all. This is where (2) has to come in. Consider the following argument introduced by Saul Kripke. In "A Puzzle About Belief", Kripke considers the case of Pierre whom we can credit with both the beliefs expressed by the sentences 'Londres est jolie' and 'London is ugly' respec-

tively. Obviously, the two sentences ascribe incompatible properties to what is in fact one and the same city. But, as Kripke shows, we can construct a coherent story as the consequence of which Pierre can be said to be disposed to assent to both those sentences at the same time. We may read this as a case in which (1) is fulfilled without there being any alteration in the subject's cognitive state at all. It would at first be true that Pierre believed of London that it was pretty and then it would be true that he believed of London that it was ugly, simply because it was all along the case that he held both beliefs. The task of (2) is thus to ensure that a real change occurs, that the subject holds the two beliefs successively.

Having said all that, why do I think that (3) is necessary? In formulating (1) and (2) I have chosen *de re* characterisations of the beliefs in question.⁶ On the one hand, this means that (1) and (2) hold independently of how the object of the thinker's belief is described (they are *referentially transparent*). On the other hand, (1) and (2) do not entail that the subject knows that the beliefs concern the same object (they are not *epistemically transparent*). As they stand, (1) and (2) leave out one fact which seems to be vital for there to have been a change of mind: namely that your believing of x that it is not- φ , *replaces* your believing of x that it is φ . That means, the latter thought-episode occupies that position in your cognitive economy which used to be occupied by the former thought-episode, and it has a different truth-value. Kripke's story shows exactly what is missing from the account as long as all we have are (1) and (2). If it is possible to hold contradicting beliefs concerning one and the same object at the same time (as part of one state of mind, as it were), holding them successively cannot be what constitutes a *change* of mind. Thus, there are cases in which (1) and (2) are true but no change of mind has occurred. These are cases in which you don't know that (1) and (2) hold.⁷ This is why (3) is necessary within an account of what it means to change one's mind. The particular problem I wish to discuss in the

6. Note that I speak of a *de re* characterisation of the beliefs, and not of *de re* beliefs.

7. As it stands, this is probably too strong, because it seems likely that all that is actually needed is your being *able* to know that (1) and (2) hold. However, that implies that at least on some occasions (3) will have to hold if a subject is to be able to change her mind. Also, the condition concerns only the immediate (temporal) vicinity of the change. I do not mean to say that you only change your mind properly if you continue to remember that you did so. What we are interested in is the dynamics of the change itself (which does to some degree involve

rest of this paper concerns the nature of the knowledge postulated in (3). In a series of papers, Ruth Millikan has recently put forward a view to the effect that, precisely since (1) and (2) can hold without (3) holding, (3) can never be the case as an *a priori* matter. The fact that the contents of two thought-episodes contradict each other (or not) cannot be transparent to the subject just in virtue of her entertaining those thought-episodes, it is something that has to be established.⁸ The subject has to re-identify the referent of one thought-episode as the referent of the other before she can arrive at a judgement about their relative truth-values.

In what follows, I shall discuss three corollaries which I take to follow directly from the claim that in order for you to change your mind you have to know that you do so. My aim is to point out that, ultimately, the notion of re-identification cannot be made to do the work Millikan would like it to do, especially as far as the temporal content of the beliefs in question is concerned, and that we either have to give up the notion of a change of mind as involving more than (1) and (2), or we have to suppose that a subject can sometimes know *a priori* whether she does in fact entertain the same thought at different times or not. In the concluding section, I wish to show that this result is compatible with some of the externalist insights which motivate Millikan's own view.

3.

The first corollary I wish to discuss concerns the kinds of attitudes you have to be able to entertain in order to change your mind:

C1: At t_1 you knew of x that you believed it to be φ , and now, at t_2 , you know of x that you believe it to be not- φ .

what happens before and after), but one can continue to hold a belief while forgetting how one acquired it.

8. This also seems to be Kaplan's view: "[I]n natural language every new syntactic occurrence of a true demonstrative requires not just a referent-determining intention, but a *new* referent-determining intention" (Kaplan 1989: 588). It would therefore seem that, for him, cognitive dynamics must come about as a mere epiphenomenon, the result of a succession of intentions which are in effect independent of one another. The question as to why these intentions occur, i.e. why there should be an obligation on the subject to adjust the expression of her thoughts to the passing of time, is strictly speaking not part of the project of direct reference semantics.

In short, a capacity to entertain higher-order attitudes is presupposed in the ability to change one's mind. Changing one's mind is a matter of possessing and exercising knowledge of one's own mental states. That this is indeed a requirement on our everyday notion of a change of mind might come out more clearly when we consider why we should find it at all worthwhile to distinguish between a change of mind and the mere acquisition of a new belief. This distinction came out in our discussion of condition (1) above, and I think where we draw such a distinction, it is because of considerations such as the following.

In her study of the phenomenon, Annette Baier locates the origin of the practice of changing one's mind in the subject's belonging to a community of "mutual regulators" (Baier 1978: 167). It is in others that we first become aware of a possible mismatch between what they believe and what is the case, and it is others who first become aware of a possible mismatch between our own beliefs and what is actually the case.⁹ And only in response to mutual criticism can a capacity for self-criticism develop which leaves room for a change of mind. However, if the adoption of a self-critical stance is the result of such an internalization of criticism made by others, this will be because the subject herself becomes aware of a possible dissociation between what she believes and what is the case. The shift from the interpersonal to the intrapersonal introduces a perspective onto one's own cognitive states. The significance of such a shift does not lie in detaching changes of mind from external factors. We often change our mind because we have found out something new about the world. But finding out something new about the world does not suffice for making us change our mind. When we do change our mind as a result of a new discovery, this is because this discovery has not just given us a new view of the world but because it has also given us a new view of ourselves, because it has fed into a self-critical faculty: into an awareness that is directed primarily onto what we believe rather than what is the case.

We don't have to agree with all aspects of Baier's speculative account of the origin of thinking that allows for changes of mind in order to acknowledge the fundamental point that lies at the heart of her considerations. No

9. Baier does not put it in those terms, but I think this is a fair translation of what she says into talk about beliefs. See also Astington & Gopnik (1988), esp. pp. 201f., for similar considerations from the point of view of developmental psychology.

matter what particular reason we have for changing our mind about something, the reason why reflection, experience, or the influence of others make us change our minds ultimately has something to do with rationality. Talk of a change of mind has its place in an explanation of the cognitive economy of a thinker.¹⁰ It ascribes to the thinker an ability to make sense of a distinction between a change in the world and a change in her own beliefs about the world, a distinction which would not make sense were it not for her ability to know about her own beliefs.

It is important to point out that C1 can be fulfilled without the subject knowing that it is in fact the same x both of her beliefs are concerned with. As Burge points out,¹¹ in order for the subject to have the type of self-knowledge required by C1, it is not necessary that she have any independent way of identifying x as the object of her first-order belief.

The source of our strong epistemic right, our justification, in our basic self-knowledge is not that we know a lot about each thought we have. It is not that we can explicate its nature and its enabling conditions. It is that we are in the position of thinking those thoughts in the second-order self-verifying way. Justification lies not in the having of supplemental background knowledge, but in the character of the self-evaluating judgements (Burge 1988: 660).

In this sense, knowing what one believes at a time does not imply knowing the object of one's belief in a way that would enable one to recognize it as the object of other beliefs. The subject can be authoritative about her own

10. In fact, it is an explanation of a skill which children have to acquire: "Three-year-olds have beliefs, that is, representations of the world, and they change those representations, but they have no understanding of representational change. That is to say, they do not know that their beliefs have changed" (Astington & Gopnik 1988: 193).

11. See also Millikan (1993: 228n), who emphasises that her claims are compatible with Burge's (1988) and Davidson's (1987) accounts of self-knowledge. It is not immediately obvious, however, how Davidson's views translate to knowledge of one's own indexical attitudes. What he says is that "[t]he agent herself [...] is not in a position to wonder whether she is regularly using her own words to apply to the right objects and events, since whatever she regularly applies them to gives her words the meaning they have and her thoughts the contents they have" (Davidson 1987: 456). It will be my aim to explore a view of indexicals exploiting similar principles, but it is clear that a slightly different story has to be told about what gives these their meanings.

beliefs in the way described by C1 even though she does not realize that they are concerned with the same object.

4.

Sometimes we can tell that a change occurs just by perceiving it occurring. In this respect, one's knowledge that one has changed one's mind must be different from one's knowledge of other kinds of changes, because it necessarily involves memory:

C2: At t_2 , you know of x that, at some time in the past, you believed it to be ϕ .

This second requirement brings in a new element: The scope of the subject's knowledge ranges no longer just over the current beliefs, but also over a belief the subject no longer holds.

Again, it is important to point out that C2 can be fulfilled together with the second half of C1 without the subject knowing that the respective beliefs are concerned with the same x . We will eventually have to ask how the subject could come to possess such knowledge, but we better not presuppose it here if we don't want to beg the question at issue. Instead, the question I wish to address is whether we can make sense of C2 without assuming that the subject already knows that the belief mentioned there concerns the same x as the belief mentioned in the *first* part of C1. The problem, in short, is that this case differs in an important respect from the cases of self-knowledge discussed above. The externalist account of self-knowledge outlined in the previous section relied on the fact that the subject entertained the first-order thought *while* ascribing it to herself, simultaneously. Only because of this fact can the reflexive judgement "simply inherit[] the content of the first-order thought" (Burge 1988: 656).¹² There seems to be no similarly elegant solution available with respect to present second-order thoughts about one's past first-order thoughts. It seems that either the subject has to know about the co-reference of her past and present beliefs, or the content of her past belief is not something she can be authoritative about.

The formulation I have chosen for C2 already entails that the subject will know of the co-reference of her past and present thoughts. In so far as C2

12. Burge's choice of words in this case seems slightly at odds with his insistence on the strict simultaneity of the two thoughts.

involves a present belief about an x , it will be transparent to the subject of that belief that this present belief concerns the same x her past belief was about. It will be objected that all this shows is that the formulation of C2 I have given begs the question. All we can demand of the subject is: 'At t_2 , you know that, at some time in the past, you believed of some x that it was φ .' This seems to defuse the problem, since, at t_2 , x is referred to under a description (roughly: 'Whatever x it was that my earlier belief was about'), which leaves open whether I can still identify it or not.¹³ If we take this reformulation as our basis, the question as to whether a change of mind has occurred or not becomes a matter of the subject's ability to re-identify x as the joint object both of her present mental state and her past mental state, an ability which is in principle fallible. This notion of 're-identification' plays a crucial role in the account Ruth Millikan gives of indexicals. At the heart of Millikan's theory is the following consideration:

To know what an indexical indexes, to identify the indexed, requires that one have a second route to thinking of it, a route other than via the indexical token, and that one grasp this second route as one arriving at the same referent. Indexicals do not in this sense tell what they point at, what it is that they bear their adapting relations to (Millikan 1993: 271).

As far as the relation between two different occurrences of indexicals is concerned, Millikan takes this to imply that the question as to whether they index the same thing or not can only be solved by investigation into what it is they in fact point at. In some cases, we are able to keep track of an individual, forming different indexicals in the process of gathering information about one thing. However, as Millikan points out, "the activity of tracking is done in the world, not in the mind" (Millikan 1993: 358). It may provide us with an ability more or less reliably to *hit upon* the same thing (the same content,¹⁴ as Millikan would say) at different times. But this doesn't make the fact that it is one and the same object the subject keeps track of an epistemic given, since there is no *a priori* guarantee against losing track

13. As Peacocke points out, this move means that we have to find a notion of 'understanding' which does not rely on sameness of thought, a notion which "marches with reference to thoughts, rather than employment of them" (Peacocke 1981: 192f.).

14. Contents, for Millikan, are what David Kaplan calls 'contents'. "In general, the result of evaluating the content of a well-formed expression α at a circumstance will be an appropriate extension for α " (Kaplan 1989: 501).

without noticing it. There is no relation between the indexicals themselves which could make their having the same object manifest to the subject. The only way the subject can identify the contents of two indexicals with each other is in terms of an identification of a different kind: by identifying each of them with the content of the same non-indexical thought which can serve to mediate between them (cf. 1993: 275 and 360).¹⁵

In my opinion, the idea of re-identification turns out to be hopeless when it comes to explaining how we could come to know what we used to believe. It is constitutive of the process of re-identification as Millikan describes it that one and the same object can be re-encountered on various occasions and that the different episodes of encountering it are independent of one another. However, precisely because there is a difference between change in the world and representational change, knowing what it was I believed cannot be a matter of being able to re-identify it with something I could find myself believing, now, independently of what I used to believe.¹⁶ What I used to believe concerned how things were at a particular *time*, and, in virtue of their very nature, times are not things we can re-encounter,¹⁷ thus, invoking the notion of re-identification here is spurious.

15. There are many aspects of Millikan's analysis with which I am in complete agreement, such as the point that keeping track does not by itself yield informative thoughts of the type '*This*₁ is identical with *this*₂' (cf. Millikan 1993: 358). Where exactly I think my account diverges from hers is the subject of the concluding section of this paper.

16. Corresponding difficulties emerge with an account of 'I'-thoughts, because we don't re-identify ourselves in different ways across different contexts. It is therefore consistent with her views that Millikan treats 'I' as an "active self name" (cf. Millikan 1993: 274) rather than an indexical. However, I cannot see how a similar solution should be available in the case of temporal expressions.

17. Elsewhere Millikan acknowledges this difficulty and suggests a solution: "It is hard to see how one could gain more than one perspective upon an affair concerning *when* this or that happened, or develop the notion of a linear time sequence for events at all, without turning to the aid of others who confirm one's judgements. Only because my judgements of time sequence nearly always match those of others do I have reason to believe that I have an objective *ability*, not merely a seeming ability, to fit things remembered into a time sequence and that objective time sequences exist at all. The medium that is another person who speaks to me provides at least the most accessible way of having more than one perspective upon time" (Millikan 1984: 309). In my view, however, her use of the notion 'perspective' here begs the question.

The point is this: if we don't allow C2 to hold in the way I formulated it, we can never know whether we have changed our minds or not, neither as an *a priori* matter nor as an *a posteriori* matter. Above, we saw that we have to restrict the domain of objects 'x' can stand for in order to make sure that we talk about a change merely in the subject's beliefs about the world rather than a change in the world itself. The beliefs in question can concern an object only in so far as this object does not undergo change with respect to the property salient for the truth of the subject's beliefs. However, the domain of objects thus described just disappears out of sight when we don't allow C2 to hold. Not only does it become impossible to retain a particular belief about what was the case at a particular time or over a period of time, once that time has passed, it becomes impossible to say what it would mean to entertain *any* belief about what was the case then. Allowing C2 to hold removes this difficulty, but it also means that there have to be cases in which we know that a present state we are in concerns the same object as a past state we were in.

5.

The main point of the above discussion can be brought out in a third corollary which also makes explicit the direct connection between changing one's mind and cognitive dynamics:

C3: At t_2 , you know of x what it would take for you to still believe that it was φ .

This corollary claims that your knowledge that you've changed your mind ultimately rests upon the fact that, had you not changed your mind, there would be modes of thinking available to you which would make this immediately transparent to you. In short, you can sometimes be directly aware of the fact that you entertain the same content on two different occasions (even though you might have to express this content in different ways).¹⁸ And your being authoritative about your changes of mind rests

18. In fact, I am putting forward an option which Millikan does not consider, because she seems to assume that on any kind of theory co-referring indexicals used at different times cannot express the same mode of presentation since one of them would have to be a 'mode directed through memory' (Millikan 1993: 338).

upon your being authoritative about holding on to the same content in this way.

This is precisely the view Ruth Millikan tries to avoid at all costs, at least if it is to express a requirement on what has to be epistemically given to the subject.

[T]he ability to represent, to think of, a kontent does not entail the ability to reidentify that kontent when one thinks of it again. Thinking of a thing does not involve 'knowing what' one is thinking of, if that means always recognizing it when encountered again in thought. This is because the identity versus difference [...] of thought kontents always depend on how matters stand in the world, not just in the head. It is never dictated by the insides of thought alone (Millikan 1993: 357f.)

Two different claims make their appearance in this short passage: there is, first, the claim that whether two thought-episodes do in fact concern the same thing or not (have the same kontent or not) depends on externalist constraints, on how the world is. Second, there is the claim, assumed to follow from this, that it cannot be part of the definition of thought that the subject should be able, upon entertaining a thought-episode, to know what it would mean to entertain a thought-episode concerning the same object at a later time.

On closer inspection, the general shape of the argument should look familiar.¹⁹ Compare the 'argument from illusion': We know that our senses sometimes deceive us, in the sense that there is no internal guarantee that how things look to us is how things actually are in the world. Some cases of things looking thus and so happen to be due to things being thus and so while others aren't. From this it is argued that, since the subject will never be able to tell for sure which of these two cases is actually fulfilled, how things actually are can, strictly speaking, never enter into our definition of what it means for things to look a certain way. Instead, we need to appeal to private 'sense-data' in order to explicate the meaning of the term.

As Paul Snowdon has pointed out, if we see ourselves forced to draw such a conclusion, it is because of a background assumption which we haven't in fact provided an argument for, namely that "there is in all cases a

19. Davidson (1987) also notes a correspondence between the problems of scepticism about the outer world and authority (or lack of it) about one's own mental states.

single sort of state of affairs whose obtaining makes 'looks'-ascriptions true" (Snowdon 1981: 203). Instead of this assumption, Snowdon suggests

a disjunctive picture which divides what makes 'looks'-ascriptions true into two classes. In cases where there is no sighting they are made true by a state of affairs intrinsically independent of surrounding objects; but in cases of sightings the truth-conferring state of affairs involves the surrounding objects (*ibid.*).

This disjunctive picture allows us to see how an instance of it appearing to the subject that such-and-such is the case can actually be counted as the subject's seeing that such-and-such is the case, without that implying that this is somehow indirectly given to the subject and that her knowledge rests upon her drawing an inference from a more 'immediate' object of acquaintance to the obtaining of the objective fact. This also means that while our entertaining certain thoughts requires circumstances outside ourselves to be a certain way, their being a certain way makes ways of thinking available to us which wouldn't be available otherwise. This is a point repeatedly emphasised by John McDowell, who points out that the disjunctive account removes the apparent epistemic barrier which seems to make our perceptual beliefs stop short of the facts. "When someone has a fact made manifest to him, the obtaining of the fact contributes to his epistemic standing on the question" (McDowell 1982: 215).

Now let us go back to Millikan's argument. As we have seen, the crucial step in her argument lies in saying that because external conditions determine whether we do in fact entertain the same content on two different occasions, considerations about sameness and difference of contents cannot enter into a definition of what it means to entertain a particular thought. We can now see that the fact that we can sometimes be mistaken about the co-reference of two of our thought-episodes does not have to stand in the way of making the fact that a thought can be entertained at different times criterial for it being the thought it is. If we can give a disjunctive account of such cases, the fact that apparent transparency can sometimes mislead us does not have to mean that we can never trade upon identity. While certain conditions outside our present mental states have to be fulfilled in order for us to be able to entertain certain thoughts, their being fulfilled can make modes of thinking available to us which wouldn't be available otherwise (as we have seen above).

The discussion of C2 was meant to show that in order to make sense of the phenomenon of changing one's mind, we have to allow for the identity

in reference of two thought-episodes entertained at different times to be an epistemic given. We could put this by saying that there is a *transtemporal identity criterion* for such thoughts. It must be possible for the subject to rely on the fact that she is entertaining the same thought again. In discussing C3, a stronger claim has emerged: The kinds of thoughts one can change one's mind about are thoughts of which one can know what it would take to entertain them on a different occasion. It is *constitutive* of the thoughts in question here that they can be entertained on different occasions. We might call this a *transtemporal distribution criterion*. In fact, the transtemporal distribution criterion is more fundamental than the transtemporal identity criterion. We cannot start off with a complete account of what it is to entertain an indexical belief at one time which leaves open the question of what it would be to entertain that belief at some other time. Rather, it is constitutive of the belief being the belief it is that it can be entertained at different times. The cognitive role of temporally indexical thoughts can only be captured in cases where they are entertained on different occasions.²⁰

6.

Above, I have said that changing one's mind, if it is to be different from a mere change in *de re* belief, must be seen as an exercise of one's rationality. I wish to conclude by contrasting the view of rationality that emerges from my discussion of what it means to change one's mind with some of Ruth Millikan's remarks about rationality. In *White Queen Psychology* she writes:

Rationality obviously pivots on content. It is contents, for example, that must not be both affirmed and denied by the rational person. To be able to draw logical inferences from content to content but only when these contents are presented in a certain way, under certain modes, is a limitation of rationality. It is not, as the rationalist tradition teaches, a relocation of rationality into some inner, purer, safer realm. To recognize the same content as the same so as to reason about it appropriately but to do this only when this content is presented again under the same semantic mode is to be capable of moments of rationality. The fewer these moments are, the less one will be rational. Multiplying modes, then, cannot possibly increase one's rationality. To be rational, one needs to reen-

20. It is here where the difference between my account and the theories offered in Perry (1979) and Lewis (1979) is probably most obvious. Their account of indexical thoughts as 'self-locating' does not acknowledge that self-location might be possible only as part of an ongoing practice.

counter-kontent in such a way that one recognizes it. Diversifying modes of presentation does not help that (Millikan 1993: 348).

The aim of this paper was to explore an option which this quotation from Millikan does not consider. In fact, there is an ambiguity in saying that kontents “must not be both affirmed and denied by the rational person.” Millikan obviously interprets this as saying that being rational *consists in* not affirming and denying the same kontent. However, this strongly externalist notion of rationality makes it mysterious how any creature could come to be such that its psychology follows rational patterns. A weaker interpretation of the same sentence suggests that rationality is that which normally *prevents us* from affirming and denying the same kontent. This does not mean lapsing into the opposite extreme of not acknowledging any externalist constraints upon thought. Indeed, the passage from Millikan is highly instructive in pinpointing where externalist constraints may come in. She is right in saying that any multiplication of modes of presentation results in a decrease in rationality. But this can also be interpreted as saying that an atomistic picture of modes of presentation—a picture according to which it is at best an accidental property of a mode of presentation that it can be entertained at different times—is hopeless, because the very point of modes of presentation is to enable a subject to entertain the same kontent on different occasions. In short, what is dependent on externalist conditions is the *individuation* of modes of presentation.²¹ Ways of thinking about a thing are ways of retaining information about it across contexts. What the world is like will have an influence on whether we succeed in doing so, but where we do, our doing so will contribute to our rationality. It is in the ability to trade upon the identity of the objects of thought, to retain one and the same belief over time, where an ability to be rational is conferred on the subject even though this “ability to be rational loops through the outside world” (Millikan 1993: 358).

The phenomenon for which Kaplan has coined the term ‘cognitive dynamics’ thus emerges as having a pivotal role in our lives as rational agents. It seems plausible that it is quite generally the case that in order to count as making a significant contribution to a thinker’s reasoning at all, a belief formed at one time must in principle be available at a later time for confirmation or revision. If I cannot form any conception of what it would

21. For similar proposals, see McDowell (1984) and Evans (1981).

take for me to change my mind in the future with respect to an assertion made now, it is not clear how such an assertion could carry any meaning. The aim of this paper was to point out that, in a case where an assertion contains an indexical expression, confirmation or revision of it at a later time cannot draw upon a re-identification *a posteriori* of what it was about, simply because what it was about may no longer be the case. If revision of a past assertion containing an indexical is to be possible at all, it must be grounded in *a priori* connections between past and present. There must be cases in which the fact that a subject is entertaining the same indexical belief on two occasions is epistemically transparent to the subject, and it is this transparency which a theory of cognitive dynamics has to explain.²²

Department of Philosophy
University of Warwick

References

- Astington, J. W., and A. Gopnik. 1988. Knowing you’ve changed your mind: Children’s understanding of representational change. In J. W. Astington, P. L. Harris, D. R. Olson, eds. *Developing Theories of Mind*. Cambridge: Cambridge University Press.
- Baier, A. 1978. Mind and Change of Mind. In P. A. French, T. Uehling, H. K. Wettstein, eds. *Midwest Studies in Philosophy IV. Studies in Metaphysics*. Minneapolis: University of Minnesota Press.
- Boghossian, P. 1994. The Transparency of Mental Content. In J. E. Tomberlin, ed. *Philosophical Perspectives 8, Logic and Language*, 33–50.
- Burge, T. 1988. Individualism and Self-Knowledge. *Journal of Philosophy* 85, 649–63.
- Davidson, D. 1987. Knowing One’s Own Mind. *Proceedings and Addresses of the American Philosophical Association* 60, 441–58.
- Dennett, D. C. 1978. How to Change Your Mind. In *Brainstorms. Philosophical Essays on Mind and Psychology*. Brighton: Harvester Press.
- Evans, G. 1981. Understanding Demonstratives. In H. Parret and J. Bouveresse, eds. *Meaning and Understanding*. Berlin: De Gruyter.

22. I am grateful to the editor of this volume and to an anonymous referee for comments.

- Kaplan, D. 1989. 'Demonstratives; and 'Afterthoughts.' In J. Almog, J. Perry, and H. Wettstein, eds. *Themes from Kaplan*. Oxford: Oxford University Press.
- Kripke, S. A. 1979. A Puzzle about Belief. In A. Margalit. *Meaning and Use*. Dordrecht: Reidel.
- Lewis, D. 1979. Attitudes *De Dicto* and *De Se*. *The Philosophical Review* 88, 513–43.
- McDowell, J. 1982. Criteria, Defeasibility, and Knowledge. *Proceedings of the British Academy* 68, 455–79. Reprinted in J. Dancy, ed. *Perceptual Knowledge*. Oxford: Oxford University Press.
- . 1984. *De re* Senses. *Philosophical Quarterly* 34, 283–94.
- Millikan, R. G. 1984. *Language, Thought and Other Biological Categories: New Foundations for Realism*. Cambridge, Mass: MIT Press.
- . 1993. *White Queen Psychology and Other Essays for Alice*. Cambridge, Mass.: MIT Press.
- Peacocke, C. 1981. Demonstrative Thought and Psychological Explanation. *Synthese* 49, 187–217.
- Perry, J. 1979. The Problem of the Essential Indexical. *Noûs* 13, 3–21.
- Quine, W. V. O. 1960. *Word and Object*. Cambridge, Mass.: MIT Press.
- Snowdon, P. 1981. Perception, Vision, and Causation. *Proceedings of the Aristotelian Society* 81, 175–92. Reprinted in J. Dancy, ed. *Perceptual Knowledge*. Oxford: Oxford University Press.

☞ TOBIES GRIMALTOS AND CARLOS J. MOYA
Belief, Content, and Cause

In this paper we intend to analyse two cases of change in belief. These cases have been designed by John Perry and William G. Lycan in order to show that the semantic properties of a subject's beliefs, such as their propositional content, cannot explain the change in that subject's behavior. Both Perry and Lycan think that the conclusion they extract from these cases can be raised to a general thesis, namely that a fairly sharp boundary must be drawn, inside the concept of belief, between semantic properties and causal powers. Perry thinks we need a distinction between belief object, or proposition believed, and belief state. It is the latter that accounts for the belief's causal role. Lycan, in turn, elaborates on Perry's distinction by asking us to distinguish between two individuating schemes for beliefs, namely the truth-conditional scheme and the computational scheme: "The truth-conditional individuating scheme is typically imposed when what concerns us are the truth-values or other semantic aspects of beliefs; the computational scheme is imposed when what we care about is causal effects" (Lycan 1988: 86). We shall contend, instead, that both cases can be accounted for with no need to create a gap between propositional content (individuated in terms of truth conditions) and causal or computational role. We shall try to show that, in both cases, the relevant change in the subject's behavior is explained by a

Research for this paper has been funded by the Spanish Government's DGICYT as part of the project PB93-1049-C03-02. Our thanks to this institution for its generous help and encouragement. An earlier version of this paper was presented at the VII SIUCC, Barcelona, May 1995. We benefited from several comments by the attendants. We are grateful to Christopher Hookway and Josep Corbi, as well as to an anonymous referee, for their useful comments on earlier drafts of this paper.