



A New Look at Personal Identity

Andy Hamilton

The Philosophical Quarterly, Vol. 45, No. 180. (Jul., 1995), pp. 332-349.

Stable URL:

<http://links.jstor.org/sici?sici=0031-8094%28199507%2945%3A180%3C332%3AANLAPI%3E2.0.CO%3B2-0>

The Philosophical Quarterly is currently published by The Philosophical Quarterly.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/philquar.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

A NEW LOOK AT PERSONAL IDENTITY

BY ANDY HAMILTON

I. THE MEMORY CRITERION AND 'IMMUNITY TO ERROR THROUGH MISIDENTIFICATION'

The account of personal identity that follows may at first sight appear to offer a 'psychological' criterion, since it lends support to Locke's claim that memory suffices for personal identity. But I am not engaged in the traditional project of providing necessary and sufficient conditions for personal identity, conditions whose specification does not presuppose identity. There are no such criteria. The essentially self-conscious ways of knowing about oneself, including memory, constitute, and do not merely furnish *evidence* for, personal identity; there is a benign circularity here. As I shall argue, the traditional distinction between 'psychological' and 'bodily' criteria is ill-founded, since, for self-conscious subjects, the body is not simply a mass of physical 'stuff', and 'bodily' criteria have an essential psychological component. Thus in other respects my account is not Lockean; his sharp distinction between the human being and the person was, I think, disastrous. It was an inspiration for the current science-fiction approach to personal identity, an approach which continues to exhibit all the philosophical insight of a Star Trek Convention.

Rejecting Locke's criterion, Butler famously argued that memory presupposes personal identity and so cannot ground it. If memory is taken as constituting a criterion in the terms of the traditional project, then Butler's circularity objection seems decisive. However, when J.S. Mill defended Locke, he was rightly cautious about the implications of the objection, maintaining that 'The phenomenon of Self and that of Memory are merely two sides of the same fact'. None the less, subsequent opinion has endeavoured to circumvent rather than embrace the circularity. Most notably, Shoemaker and Parfit embarked on a reductionist vindication of psychological

criteria, invoking a notion of *quasi*-memory ('*q*-memory') to avoid the alleged circularity.¹

Shoemaker and Parfit define mere *q*-memories as phenomenologically identical to and caused in an appropriately similar way to ordinary memories, but originating in the experience of a person other than the remembering subject. (I shall later question the terms of this definition. Henceforth, '*q*-memory' = 'mere *q*-memory', i.e., it is assumed that the *q*-memory comes from a different subject.) Their strategy has the consequence that when I seem to remember an experience, it is merely an *assumption* I make that I am identical with the person who had the experience, an assumption justified by the contingent non-prevalence of *q*-memory.² The presupposition here is that 'I' effects an essentially present-tense identification.

This presupposition seems to be part of the traditional approach to personal identity – an approach shared by Parfit, even if he denies that the concept of personal identity has application. Hence one asks what makes it the case that person *X* at time *t*₁ and person *X* at time *t*₂ are the same person, a question one can pose in the first person only if 'I' does effect merely a present-tense identification. Gareth Evans explicitly rejected this presupposition, claiming instead that the identification 'spans past and present'. Interestingly, he continued:

This [feature] has a connection with the idea that memory suffices for personal identity. My judgement (or seeming judgement) 'I was F', made on the basis of apparent memory, can go wrong only in certain ways. Either there is no past state of affairs of which the apparent memory constitutes information (no one was – relevantly – F); or the object involved in the state of affairs was indeed myself; or finally, *I* do not exist.³

Evans' claim is that memory-judgements are 'immune to error through misidentification' [henceforth 'IEM'; acronyms are an affliction in contemporary philosophy, but I hope this one may be excused]. Proponents of *q*-memory, in contrast, maintain that IEM is not preserved in *q*-memory

¹ J.S. Mill, 'Notes to James Mill's *Analysis of the Phenomena of the Human Mind*', in *Collected Works* Vol. xxxi, 'Miscellaneous Writings' (Toronto UP, 1989), pp. 212–13; D. Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984), ch. 11, especially pp. 219–22; S. Shoemaker, 'Persons and Their Pasts', in his *Identity, Cause and Mind* (Cambridge UP, 1984), especially pp. 41–3. On Mill's view, see my 'Phenomenalism and the Self', in J. Skorupski (ed.), *The Cambridge Companion to Mill* (Cambridge UP, forthcoming 1996).

² See, e.g., D. Parfit, 'Personal Identity', *Philosophical Review*, 80 (1971), p. 15.

³ *The Varieties of Reference* (Oxford: Clarendon Press, 1982), p. 246 and footnote; most subsequent references in the text are to this volume. The final possibility is the brain in a vat scenario, which I shall ignore. J. McDowell defends Evans' line in 'Reductionism and the First Person', a paper presented in 1991 and still to appear in the long-forthcoming J. Dancy (ed.), *Parfit and his Critics* (Oxford: Basil Blackwell).

cases.⁴ Evans connects IEM with the memory criterion, without perhaps seeing the eventual consequence – abandonment of the traditional approach to personal identity. His claims are, with some qualifications, important and correct. Much of the rest of this article is concerned with their vindication, partly against arguments from a Parfitian perspective, partly against other arguments concerning *q*-memory that Evans himself presents, apparently unaware of the resulting tension.

In illustration of IEM, consider the judgement ‘I once sold Samsonite suitcases in Barkers of Kensington’. If this expresses a *memory*-judgement, I cannot then be mistaken about *who* sold Samsonite suitcases. (The Samsonite brand was a best-selling line in Barkers’ travel goods department of the 1970s; or so I seem to recall.) If, for whatever reason, I come to doubt that I had that job in that department store, it will make no sense for me to continue to maintain, *with the same justification*, that none the less *someone* used to sell Samsonite suitcases. That is, I cannot say ‘Well, I distinctly remember someone selling them’. I therefore want to insist on the dilemma ‘Either I personally remember, or it is a delusion’, thus ruling out the possibility of *q*-memory.

When, in contrast, the claim is made with a different justification, *viz.*, testimony – a friend recalls seeing me behind the counter, purveying some expensive but highly durable items of Samsonite luggage – then there is no IEM guarantee. If that friend later confessed that he could not remember whether it was myself or someone else who had been behind the counter, it would *not* be senseless for me to ask ‘Someone certainly sold Samsonite suitcases, but was it myself?’. Judgements made on the basis of testimony are not IEM, even when as in the present example they are tokens of the very same type as a memory-judgement.

I shall therefore argue that

1. Memory-judgements are IEM.
2. It follows that the remembering subject must be identical with the remembered subject; that is, *q*-memory is unintelligible.
3. There is no unacceptable circularity in (2); traditional conceptions of personal identity are mistaken.

II. MEMORY-BASED IEM: ‘APPARENT MEMORY’

Like Evans, I believe IEM is exhibited by a wide range of first-person ascriptions including sensation, perceptual experience, belief, intention and

⁴ See for instance H. Noonan’s pellucid *Personal Identity* (London: Routledge, 1989), p. 178, which is invaluable in charting the debate.

varieties of bodily ascription. But my definition of IEM is rather different from his:

An assertion of 'I am F', 'I ϕ ed', 'I shall ϕ ', etc., is IEM *if and only if* the justification subjects would offer for it is such that if they subsequently come reasonably to doubt the assertion, no matter what the cause, it will be senseless for them to cite the original justification as a reason for claiming that none the less *someone* is F (or ϕ ed, etc.).

The senselessness of the doubt constitutes IEM and is not, as Evans seems to think, merely a manifestation of it. It is essential that the 'original justification' is that which the subject would offer if asked to justify the first-person utterance (obviously the subject's sincerity is assumed). IEM may thus be regarded as an 'internalist' and not an 'externalist' or causal notion, since it concerns what the subject is justified in believing.⁵ In emphasizing that 'no matter what the cause' of the doubt, the retreat claim 'Someone is F' makes no sense, the purely contingent immunity that may occur in cases of testimony, etc., is ruled out. For example, I may come to believe on the basis of testimony that I had a serious accident when a child, but then discover that my informant is totally unreliable. I would then have no reason to believe that *anyone* had an accident, since the justification for the judgement is totally destroyed by this later information; but this would not be a case of IEM in the relevant sense.

What is the 'original justification' in the case of personal memory-judgements? (We are not, of course, concerned with factual, 'learned and have not forgotten' memory.) A past-tense claim constitutes a personal memory-judgement only when the subject is prepared to offer the justification 'I remember ϕ ing' or 'I seem to remember ϕ ing'. That is, if someone asks how I know that I sold Samsonite suitcases, I reply, perhaps sarcastically, 'I *remember* selling them!'; or alternatively and tentatively, 'I seem to remember selling them'. This formulation applies to witnessings as well as actions: if someone asks how I know that my colleague was delighted that his article had finally been accepted by *The Philosophical Quarterly*, I may reply 'I remember seeing his reaction when he got the letter of acceptance'. (If it was a recent event, I would say simply 'I saw his reaction'; but I shall ignore complications over short-term memory.)

⁵ To use the fashionable distinction, which is better conceived in epistemology than it is in philosophy of mind, where it has greatly exercised the Star Trek tendency. For Evans' account of IEM, see pp. 179–91, 215–25, 240–9. He uses a variety of locutions concerning the relativity of IEM to a 'ground', as he puts it. (I prefer 'justification', which is more neutral as to whether memory is a source of new, or merely retained, knowledge.) Evans claims that IEM is relative to a 'way of knowing' about oneself (p. 218), but also that it is relative to 'the subject's having ... or appearing to have information' about himself (p. 221, see also pp. 242 and 245); I shall question the latter formulation below.

The justification requires qualification, however. Goethe says in his autobiography that he is going to tell of things belonging to his childhood, of which he does not know whether he remembers them or was told them. Note the correctness of his description. If he had been told, this would not be a case of *remembering* in the personal sense, even though it involved 'habitual' memory.⁶ An example would be my own recollection of swallowing a ball-bearing when a young child, and being taken to hospital. As far as I know, this event actually happened. But it might have been something I had been told; perhaps it seems to be a personal memory, but is not. Say my mother denies it happened to me, but has a vague recollection of its happening to someone. It would then make sense for me to say 'Perhaps I was told this story. So maybe it did not in fact happen to me, perhaps it was a friend of mine who swallowed the ball-bearing.'

This example shows that Evans is wrong to characterize the basis for IEM simply as 'apparent memory'. 'I seem to remember swallowing a ball-bearing' implies either 'but maybe I did not' or 'but maybe I was told'. There are two kinds of 'apparent memory': 'I'm not sure if I ϕ ed' – maybe nothing of the sort happened, the apparent memory is delusive – and 'I ϕ ed, but I'm not sure whether I genuinely remember ϕ ing or was told that I had ϕ ed'. (For the second kind, Anscombe has 'I at least ostensibly remember', which she curiously regards as ordinary usage; I suppose it may be so in Cambridge.) In the latter situation, it makes sense for the subject to say 'Well, I at least remember that *someone* swallowed a ball-bearing'. But this is not *personal* remembering. In retreating to this position I come to believe that the impressions, if any, that I took to be memory-impressions of swallowing, distress, being taken to hospital, etc., are delusory. In so far as I entertain the retreat claim, I am wondering 'Do I really remember this?', i.e., perhaps my 'recollection' is based on testimony. 'I remember ϕ ing', rather than 'I remember that I ϕ ed', is the justification characteristic of personal memory. 'I remember that I had a marvellous breakfast' when applied to events of ten years ago is at best merely consistent with personal remembering. (Much more could be said on the problematic distinction between factual memory of my own past and personal memory.)

⁶ On factual memory and the Goethe example, see G.E.M. Anscombe, 'Memory, "Experience" and Causation', in her *Collected Philosophical Papers*, Vol. II (Oxford: Basil Blackwell, 1981), p. 120. Such cases are standardly taken to support a causal analysis of memory, on the grounds that they show the need for an appropriate causal connection, 'from the inside', between past event and present apparent recollection – see for example R. Wollheim, 'Memory, Experiential Memory and Personal Identity', in G. Macdonald (ed.), *Perception and Identity: Essays Presented to A.J. Ayer* (London: Macmillan, 1979). For reasons suggested below, I do not favour a causal analysis of memory. See also Anscombe *op. cit.*; and R. Squires, 'Memory Unchained', *Philosophical Review*, 78 (1969), pp. 178–96, who argues that memory involves merely the retention of knowledge, and that retention is not a causal notion.

There will be a suspicion that if 'I seem to remember ϕ ing' is the original justification, then the existence of 'apparent memory' based on testimony shows that there is no IEM guarantee *at all*. But it is important to note that the subject's inclination to say 'I seem to remember ϕ ing' is not a complete description of the situation. Such utterances are not reports of an introspectable state of 'seeming', but rather, tentative claims motivated by a doubt – and the kind of doubt must be specifiable. If I wonder whether the event really happened, then the subsequent appearance of a doubt about whether I was told will constitute a *new* justification in the terms of the IEM definition. So a full specification of the 'original justification' is as follows: the subject is prepared to say 'I remember ϕ ing', or 'I seem to remember ϕ ing', where the implied doubt is 'but maybe I did not'.

Recognition of 'apparent memory' based on testimony has two important consequences. First, subjects have no infallible knowledge of their identity over time. There is no introspectable state, known with certainty, and common to all cases of sincere, spontaneous past-tense claims, since there is no *introspectable* difference between the two cases of 'apparent memory'. (This contrasts with present-tense cases where sensation grounds IEM but visual perception does not; thus it makes no sense to say 'My hair is blowing in the wind, but I am not sure whether I feel this or see it'.) In both cases of 'apparent memory', mental images may be present, though the attitude towards them will differ; Goethe would not have doubted that his images had some relation to past events. The denial of infallible knowledge of identity becomes important when the circularity objection is reconsidered in the final section below.

A second consequence is that there is something akin to *q*-memory in normal experience. Evans does allow that subjects *q*-remember an event if and only if (i) they have an apparent memory of such an event, and (ii) that apparent memory embodies information deriving from the perception of that event by perceivers who were not necessarily themselves. Since this definition actually fits 'apparent memory' based on testimony, it is inadequate; and, I shall argue, it cannot be supplemented to produce any kind of personal memory. Testimony-based 'apparent memory' is precisely *not* personal remembering. Memory-judgements are spontaneous; i.e., not made as a result of receiving indications after the event, whether statements or evidence, that the thing happened.⁷ This suggests that so-called '*q*-memory', where it is construed as involving indications after the event that in fact it happened to someone else, cannot be a kind of personal memory, as I shall now argue.

⁷ This goes somewhat beyond Anscombe's definition of 'spontaneous' (*op. cit.* pp. 124–5).

III. *Q*-MEMORY

Q-memory is standardly defined in terms of the kind of causal or functionalist account of memory I wish to reject, one that regards memory as an 'informational state' occupying an appropriate causal role, involving an 'internal' 'information-link' between past experience or witnessing and present recollection. But I shall go along with these standard definitions initially. What I shall call '*strong q-memory*' is, as Evans puts it, the faculty whose 'informational states' 'announce themselves ... as *merely q*-memories, so that it seems to the subject that someone or other ϕ ed without its in any way seeming to him that *he* ϕ ed' (p. 248). As Evans argues, such a faculty is possible only if Parfit is right to say that 'it cannot be a part of what I seem to remember about [an] experience that I, the person who now seems to remember it, am the person who had this experience.... This is something I am justified in assuming only because I do not in fact have *q*-memories of other people's experiences'.⁸ Parfit's claim presupposes that 'I' effects an essentially present-tense identification.

'Strong *q*-memory' clearly runs counter to memory-based IEM. It licenses claims such as 'I seem to remember someone having a holiday in Bournemouth, though I am quite unsure whether it was myself who had it'. Both Evans and McDowell rightly find this quite unintelligible. I agree with them that 'strong *q*-memory' cannot be a faculty of memory; its deliverances are not spontaneous past-tense claims about events speakers have witnessed or experiences they have had.

Unlike Evans and McDowell, however, I find the alternative, 'weak *q*-memory', only superficially more plausible. This is because if *any* faculty of *q*-memory is to have wider appeal than its materialist or functionalist origins, it must finally license past-tense claims of the sort, as in the Bournemouth claim, which Evans and McDowell find objectionable. But I anticipate. As noted above, Evans' bipartite definition of '*weak q-memory*' is satisfied by 'apparent memory' based on testimony, so a third condition is required, as follows: subjects (weakly) *q*-remember an event if and only if

1. They have an apparent memory of such an event.
2. That apparent memory in fact embodies information deriving from the perception of that event by a person who may not be identical with the rememberer.

⁸ Parfit p. 15; Evans p. 246. This is precisely the sort of *q*-memory advocated by Noonan, pp. 182–5.

3. The subjects were not told or did not otherwise receive the information in a non-memory-like or 'extraneous' way. (That there should be an 'information-link' 'from the inside' would be the causal analysis, which for reasons to be indicated I want to reject.)

Reference to 'apparent memory' is rendered unproblematic by clause (3). An example of 'weak *q*-memory' would be: I seem to remember going on holiday to Bournemouth (unlike 'strong *q*-memory', there is here no initial doubt about *who* went on holiday). It turns out I never had a holiday there, but the apparent recollection, instead of being a delusion, in fact 'embodies information' from someone else's holiday in Bournemouth. So I merely '*q*-remember' going on holiday there. Evans allows that 'weak *q*-memory' in this sense is possible, but denies that it could constitute *knowledge* of, as opposed to information derived from, the past of someone who was not myself (p. 244); or at least, he denies that if it served as the basis of knowledge of that past, this could be a normal operation of memory (p. 245). But it is not really clear what he thinks – as we shall see when we return to the question of circularity. I shall first give an argument that constrains *q*-memory, however conceived, then give two sets of arguments against 'weak' *q*-memory.

Argument 1: distinctness of subjects not a datum

The prevalence of *q*-memory must be limited, since distinctness of subjects is an *interpretation* and not a *datum* of the *q*-memory scenario. A case must be made for saying that the subject who had the experience and the subject who later remembered (or '*q*-remembered') it are distinct, as against saying that they are identical. There is a question whether the correct description is '*Y* has *q*-memories originating in *X*', rather than '*X* and *Y* are the same subject (one "phenomenal" body, two "physical" bodies)'. Writers such as Noonan neglect the second possibility.⁹ Which is the correct description is not something that is necessarily decidable prior to the bizarre events implied, even assuming they are conceivable at all, which I shall shortly be questioning.

It must surely be conceded on all sides that there is *some* limit of this sort; the concern that, in pushing it hard, I am *assuming* a 'psychological' criterion will be resolved below. Materialists or functionalists, in assuming that the presence of some information-links need not affect the identities of subjects, will argue that there could be *limited* occurrence of *q*-memories between distinct subjects. So I shall first focus on these materialist or functionalist assumptions, and then see if limited *q*-memory is possible without them.

⁹ Noonan writes (p. 191) of 'other possible [*q*-memory prevalent] worlds, to which our concept of a person still has unproblematic application ...' – has he visited them to check?

Argument 2: functionalism and 'information-links'

It is, standardly, materialist or functionalist assumptions (the latter as specified above) which underwrite the identification of some spontaneous but erroneous reports, normally regarded as delusions, as a 'memory-like' phenomenon yielding knowledge of someone else's past. Without these assumptions, Evans' talk of person-duplication (p. 241) and memory-transplants (p. 244) makes little sense. But the supposition of a determinate inner process, characterized in functionalist terms as an 'information-link', should be resisted. It persists, perhaps, due to the following natural picture. I observe that it is raining. A week later I recall that it was raining last Monday. Information must have been stored in me, otherwise I could not have recalled that it was raining. I have the memory 'It was raining last Monday', and that memory is constituted by the stored information.

But a stored inscription is at best a precondition of personal remembering. There must be the further understanding that the remembering subject is inclined spontaneously to assert the inscribed sentence. The inner 'informational state' is at best derivative. This claim is perhaps supported by Ayer's laconic refutation of causal analyses:

Past events cannot be introduced as the otherwise unknown causes of our apparent memories of them. We cannot pick and choose among them in order to forge the right causal links, unless they have been independently identified. And how can this be achieved except through memory?¹⁰

His point may be illustrated by examination of 'information-garbling', which seems to offer a counter-example to IEM. Say I reminisce about a 'brilliant objection' I made in a particular philosophical discussion, but my audience, non-plussed, do not recall any occasion on which I made that point. But then someone recalls Brennan making it in the discussion; so my information was 'garbled'. But note that the evidence from testimony is a *new* justification, leaving unaffected the IEM status of the original judgement, which should be seen as a delusion. A determinate 'information-link' from the discussion to the mistaken reminiscence is *assumed* by talk of 'garbling'; but in practice we have to start with the datum of the 'reminiscence' and try to explain it without such an assumption. 'Information-garbling' is just a metaphor, and a pernicious one; so is the concept of an 'information-link' from which it is derived.

These compressed claims may be developed by showing that the mere postulation of 'information-links' between subjects will not generate *q*-

¹⁰ A.J. Ayer, 'The Causal Theory of Perception', in his *Freedom and Morality and Other Essays* (Oxford UP, 1984), p. 81; I would not in general want to endorse his account. I am also indebted to arguments concerning belief in A. Collins' *The Nature of Mental Things* (Univ. of Notre Dame Press, 1987), ch. 2.

memory. Evans, given his functionalist approach, sees no problem with the very idea of an 'information-link' that is not recognized as a way of knowing about an object. But both he and McDowell argue against Parfit that the subject is able to distinguish only between genuine memories and delusions (they would also, I presume, allow 'apparent memory' based on testimony). Of course if such an information-link came to be recognized as reliable, Evans believes, I would have at best an *inferential* and hence *non-memory* justification for 'I had a holiday in Bournemouth', since 'strong *q*-memory' is incoherent. He and McDowell seem all the same to agree with Parfit that, from a superior scientific psychological vantage point, those delusions *in fact* constitute (weak) *q*-memories – defined as 'information' originating in the past of someone other than the rememberer. This view is connected with Evans' apparent espousal of Butler's circularity objection, as we shall see.

But there is no such superior vantage point; the Evans–Parfit dispute should be undercut. Information is not an 'impersonal' notion; it does not just physically lie around in the universe waiting to be gained. Information is only information *for someone*, i.e., when it is recognized as derived from a way of knowing. Tree-rings do not constitute 'information' about the age of trees until we learn to calculate it by using them. In considering *q*-memory, the data with which we have to deal are the spontaneous past-tense reports of subjects, plus any correlations that may be drawn between (by normal standards) erroneous reports and the past of an individual distinct from the maker of the report. 'There is an information-link' means nothing more than 'People infer from such phenomena to facts about the past of another, and reliably so'. Until these correlations lead us to postulate a way of knowing, there is no (unrecognized) 'information-link'.

Argument 3: q-memory without materialism or functionalism

Maybe *q*-memory does not require questionable assumptions of functionalism, 'information-links' and superior vantage points. Perhaps I just have to see that my putative memory-claims fit someone else's past. This is the most conservative presentation of *q*-memory, and it yields the strongest case for it.

Let us develop the curious case of the Bournemouth holiday. Reminiscing about the holidays of my childhood, I seem to remember arriving by steam train at Bournemouth station, armed with bucket and spade, etc. I recall the busy platform, the station nameboard, the noise and smell of the steam-engine. However, my mother assures me that we never had a family holiday in Bournemouth, and in any case, so I am told, steam trains ceased operating that service before I was born. But my reminiscence has caught the ear of fellow philosopher D. Parfit, and the details I impart remind him uncannily of a childhood holiday *he* had in Bournemouth. I draw, apparently from

memory, a picture of the hotel where I thought I had stayed; Parfit comments on the amazing likeness to the one he stayed at. And so on.

Proponents of *q*-memory will argue that, seeing the mounting evidence of a 'fit' between my apparent memory-judgements and Parfit's past, I shall at some point be persuaded to postulate a '*q*-memory-link'. But there is no appropriately memory-like justification for doing so. 'I at least remember *that* Parfit arrived' is not sufficient for personal remembering. 'I at least remember Parfit arriving at Bournemouth station', or 'I at least remember someone arriving' both fail to capture the experience recalled, since as far as the content of what is recalled is concerned, Parfit, or a mere 'someone', did not come into it. One could try 'I remember arriving at Bournemouth station – only the "I" then was Parfit!'. (It was the same 'I', only the person had changed his identity.) Or 'I remember arriving at Bournemouth, only it was not my body that arrived'. But these are desperate expedients. As Wittgenstein said of Mr Ballard, this is 'a queer memory phenomenon'.¹¹

'I seem to remember someone having a holiday in Bournemouth, though I am quite unsure whether it was myself or someone else' is an expression of profound puzzlement – not a report of an introspectable state, or a final deliverance of a memory-like faculty. I may conceivably come to *infer* from 'I seem to remember arriving at Bournemouth', together with something I have learnt about my Bournemouth 'memories', to 'Parfit arrived at Bournemouth'. Whether there is an 'information-link' between myself and the past of another depends on what view people come to take about such curious (and necessarily limited) incidents. But, as argued above, their conclusion would not track any prior fact of the matter. The result is in any case not a personal memory-like faculty; the idea that genuine memories are a *sub-class* of '*q*-memories', in a way that avoids circularity in the memory criterion, is quite absurd. Development of this kind of example, rather than discussion of teletransporters and the like, is the only sensible basis for Parfitian *q*-memory; but it is not enough.

Like Evans and McDowell, I want to insist on the dilemma 'Either I personally remember or it is a delusion'. But I also want to get away from the picture of personal remembering as 'inner state + information-link (possibly to someone else's past)'; I am not sure those writers, in their tolerance of 'weak' *q*-memory, do avoid it. As a result, in an apparent revival of Butler's circularity objection, Evans seems to deny that the remembering subject and the remembered subject must be identical. That is, he rejects the idea of a necessary link between IEM and personal identity – a link that I shall now defend.

¹¹ L. Wittgenstein, *Philosophical Investigations* (Oxford: Basil Blackwell, 1958), §342. See C. Rovane, 'Branching Self-Consciousness', *Philosophical Review*, 99 (1990), p. 368, for a further expedient.

IV. CIRCULARITY, BODILY SELF-AScription AND THE UNITY OF CONSCIOUSNESS

Evans' discussion of circularity precedes his treatment of IEM, but has an undoubted bearing on it. He begins by warning of 'a danger inherent in all our reflections on self-consciousness', arguing that apparent guarantees of re-identification of the subject over time may be 'tautological'. Simply in virtue of using personal pronouns, we are 'building the subject's identity over time into the description of his situation. This may make it appear that he has an infallible knowledge of what is involved in this identity; but the appearance is nothing but an artefact of our way of describing the situation' (p. 213). Evans gives two examples of the resulting error. The first is Anscombe's guarantee of 'no unnoticed substitution' of the referent of 'I'.¹² In the second example Evans, while not apparently intending to undermine memory-based IEM, denies that the remembering subject and the remembered subject must be identical. The example is under-described and rather unclear (pp. 214–15):

We might say: certainly, when time *t* comes, [the subject] will know whether or not the hypothesis that he expressed earlier by 'I'll be in pain at *t*' was or was not correct, just [on the basis of] whether or not he is in pain at *t*.... It is not possible for the subject to have got hold of the wrong person at time *t*.

Evans does not want to deny that there is criterionless self-ascription of anticipated properties, which he regards as 'simply the other side of the same coin as criterionless self-ascription of remembered properties' (p. 215n). He continues:

Of course it is not possible for the subject to have got hold of the wrong person – as the case is described, there is a logical guarantee of adequacy. But this is, again, an artefact of our way of describing the situation.... The 'method of verification' has a presupposition. Of course we must not say (using the ordinary vocabulary): it presupposes that the subject remains the same over time. But it presupposes that the subject who exists at *t* and 'remembers' the hypothesis expressed earlier is the person who made the hypothesis, and hence is the person whom it concerns. And this is something of which he can have no genuine logical guarantee.

Evans is right to reject infallible knowledge, of course; it may turn out that I was not genuinely remembering. But since *q*-memory is incoherent, if I *am* genuinely remembering, there *is* a 'logical guarantee' of identity. This is something Evans seems unwilling to concede. Perhaps he denies the

¹² See her 'The First Person', in her *Collected Philosophical Papers* Vol. II (Oxford: Basil Blackwell, 1981).

guarantee because he regards IEM, though fundamental to self-consciousness, as consistent with the possibility of weak *q*-memory. Despite his vigorous and impressive defence of IEM, his comments here and elsewhere betray a residual and conflicting commitment to the traditional conception of personal identity. He seems to agree with Reid that, on a Lockean account, 'personal identity is confounded with evidence we have of our personal identity'. He is thus driven to the hybrid claim that self-consciousness involves ways of knowing that *normally* exhibit IEM, leaving open the possibility that abnormally they may at least 'yield information' about a person distinct from the subject.¹³

Evans should instead have acknowledged that the modes of self-knowledge which are IEM *necessarily* yield knowledge of, and indeed information from, the subject and not someone else. This leads naturally to the recognition that personal identity is not a 'further fact' 'disclosed' by memory; that the self-conscious 'ways of knowing' are not merely *evidence* for personal identity. Evans is reluctant to take this path because of his robust sense of personal identity, and his conviction that, in using 'I', the 'I'-user is tracking an object (a living human being). In contrast, I want to insist that memory neither discloses personal identity, nor constitutes it in the sense of providing a criterion that is specifiable without circularity. This is a false dichotomy, since there is a *benign* circularity in concepts; a circularity illustrated by Argument 1 above, which claimed that distinctness of subjects is not a datum in *q*-memory cases. Mill was therefore correct to claim that 'The phenomenon of Self and that of Memory are merely two sides of the same fact'.

It follows that 'self-identity' is not a 'mere artefact' of our way of describing the situation; there is no other way in which 'the situation' can be described, as Evans perhaps admits when he says that the use of personal pronouns is 'unavoidable'. There is no superior standpoint from which this 'circularity' can be eliminated. The 'guarantee' rests not on infallible knowledge, but on the arguments against *q*-memory given above, which as we shall see may be generalized to other putative *q*-phenomena. Hence my rejection of any 'further fact', constituting personal identity over and above the evidence for it, has quite different implications from Parfit's rejection of it.

The claim of 'benign circularity' may be clarified and defended by examining its historical antecedents, beginning with Locke. In his formu-

¹³ I owe this point to Lucy O'Brien, though she argues that the identity conditions of a person are not constituted by these ways of knowing alone (see her 'The Problem of Self-Identification', *Proceedings of the Aristotelian Society*, forthcoming 1994-5). See also Reid, 'Essays on the Intellectual Powers', in *The Works of Thomas Reid*, ed. Sir W. Hamilton (Edinburgh: James Thin, 1895), p. 351.

lation of the circularity objection, Butler (and perhaps Leibniz before him) assumes that the unity of consciousness consists in consciousness of an independently unitary thing – a simple substance. But this is just what Locke was questioning. And, as Michael Ayers has noted, especially after Hume and Kant, ‘it can hardly be considered beyond all question that a unitary self is prior, whether epistemologically or ontologically, to the unity of consciousness’.¹⁴ Though he would have been perturbed at the suggestion of affinity with the Transcendental Philosophy, this is the position to which Mill’s more tolerant attitude towards circularity should lead.

On Kant’s view, self-consciousness (in the sense of ‘apperception’) is not something attached to each perception, involving an external relation to a body or Ego, nor is it produced by an internal causal relation between perceptions. Rather it is a necessary unity, arising directly from the very combination or synthesis of representations; the subject is necessarily active and self-creating. This at least is the picture presented in the Transcendental Deduction (*Critique of Pure Reason*, especially B131–6). Kant is also committed to the transcendental idealist notion of the subject as object of inner sense. This notion – hopelessly obscure even by Kant’s own exacting standards – may be behind the passage from the Paralogisms that Evans cites in support of his diagnosis of circularity: ‘The identity of the consciousness of myself at different times is therefore only a formal condition of my thoughts and their coherence, and in no way proves the numerical identity of my subject. Despite the logical identity of the “I”, such a change may have occurred in it as does not allow of the retention of its identity’ (A363). The first sentence rejects the Cartesian ‘I think’ as purely formal, as not indicating a substantial self distinct from the ways of knowing. In so far as the second sentence contrasts the ‘identity of the consciousness of myself at different times’ with the genuine identity of a soul or self, it involves an unfortunate appeal to transcendent metaphysics. The ‘numerical identity of my subject’, in the terms in which Evans is committed to it, is a fiction.

Memory is one aspect, though a central one, of the Kantian ‘unity of consciousness’. Someone witnesses an event, then later recollects it. What makes it the case that these experiences are had by the same person? This is one obvious example of the ‘unity question’, which may be posed, possibly tendentiously, as ‘What is it that unites people’s present experiences with their past ones?’. Of course in the case as just described, it is presupposed that it is the same person, and so there seems to be nothing further to say. But this is more than just the ‘linguistic artefact’ that Evans suggests. Let us try to say something further, and see what happens.

¹⁴ M. Ayers, *Locke: Epistemology and Ontology* (London: Routledge, 1991), Vol. II p. 270; on Leibniz, see p. 269.

The obvious 'further fact' in virtue of which the experiences are had by the same person is suggested by Ayer: 'a person's ownership of states of consciousness consists in his standing in a special causal relation to the body by which he is identified'.¹⁵ So the unity question is answered by giving a particular answer to the identity question – by specifying bodily criteria for personal identity. Indeed it may be felt that the unity question *must* be collapsed into the identity question, otherwise it will lead to Humean absurdity; taking it in any prior or independent sense means neglecting the fact of the subject's embodiment.

Ayer's claim that bodily identity underlies the 'unity of consciousness' is none the less mistaken. Clearly consciousness is in some sense causally dependent on the functioning of a human organism; and the criteria for the ascription of conscious states make essential reference to the behaviour of an embodied subject. But this relation of dependence will not determine the unity and identity of the person, since the body cannot be defined independently of self-consciousness. The unity- and identity-determining relation is not causal, since '*my* body' cannot be individuated without reference to my self-conscious awareness of it. It is at this point that IEM helps to elucidate the unity of consciousness. It does this partly because, as Evans emphasizes, it is a feature of 'bodily' as well as 'psychological' self-ascription. IEM will thus enable us to avoid the neglect of the body that has been a feature of 'unity of consciousness' accounts of personal identity. (It is because of its role in bodily self-ascription that Evans views IEM as an antidote to Cartesianism.)

The guarantee of identity of remembered and remembering subjects is 'tautological', as Evans puts it, only if one believes that personal identity consists at least partly in some 'further fact' involving third-person criteria of 'bodily continuity'. So how does one respond to the claim that bodily distinctness itself contributes to determining personal distinctness? If one is to take seriously the claim that persons are essentially embodied, should not the criteria for personal identity have something in common with those for non-selfconscious living creatures and physical objects? Hence an objection that could have been raised against my first argument against *q*-memory, that distinctness of subjects is an interpretation, and not a datum, only if 'psychological' criteria of personal identity are presupposed. My response to this objection is to argue (i) that 'bodily' criteria themselves have an essential 'psychological' aspect, and (ii) that bodily criteria thus 'psychologized' (standardly third-person) do not provide a more fundamental basis for judgements of personal identity than that offered by traditionally psychological criteria (standardly first-person).

¹⁵ A.J. Ayer, *The Concept of a Person and Other Essays* (London: Macmillan, 1963), p. 116.

On the first issue, Strawson and Evans are clearly right to emphasize the importance of bodily self-ascription, neglected by Kant. But criteria of bodily identity and individuation are not merely third-person. Just as the concepts of personal identity and self-conscious ways of knowing interlock, so do the concepts of bodily identity and self-conscious bodily awareness and control, *viz.*, kinaesthetic sensation, awareness of posture, control of one's limbs, etc. The boundaries of my body are the boundaries of the object, judgements I make concerning which are IEM. (This involves a sufficient, not a necessary condition; the boundaries of my body do not necessarily shrink as my bodily awareness does, for instance in Anscombe's Sensory Deprivation Tank.¹⁶) So the human body is not simply a physical object; 'my body' is the body of which I have fully self-conscious knowledge and control, i.e., knowledge and control which exhibit IEM. One can say that two experiences are had by the same person if they are connected with the same body – that person's body. But talk of 'that person's body' does not pick out an object independently of that person's awareness of it.

To pursue these claims a little: Evans imagines the possibility of a 'deviant' causal chain linking the subject's brain with someone else's body, so that, as he puts it, the subject 'registers information' from that other body. This scenario, he writes, 'merely shows the possibility of an error'; I may mistakenly believe *my* legs are bent, when I am receiving information from the body of someone whose legs are bent. But 'it does not show that ordinary judgements of [IEM-based bodily self-ascription] are identification-dependent' (p. 221). However, Evans grants too much. As with memory, the alleged fact that the proprioceptive sensations originate in distinct subjects is an interpretation rather than a datum of the scenario. The extension of IEM-based knowledge and control into the body of 'another' would raise a question about the distinctness of the bodies, and hence of the persons, in question.

Again, of course, we are presented with the issue of a *limited q*-phenomenon – limited '*q*-proprioception'. At least for the duration of the bizarre episode, could not the other subject's body be regarded as an extension of mine? But the situation is in fact quite analogous to that of *q*-memory. We have hardly the remotest idea how such an 'information-link' between distinct subjects could be set up. Achieving it by the extension of a nerve-ending is only superficially easier to envisage than the case of a *q*-memory-link; and the same objections based on spontaneity of self-ascription arise. As with my 'Bournemouth' memories, I could perhaps learn reliably to *infer*, on the basis of having a certain kind of bodily sensation, to

¹⁶ See her 'The First Person'; and my 'Anscombian and Cartesian Scepticism', *The Philosophical Quarterly*, 41 (1991), pp. 39–54. Of course this qualification needs developing.

the state of someone else's body. But this would not be proprioception. And unless I can do even this, talk of an 'information-link' is premature.

So causal dependence on a body, defined in non-psychological terms, is not an adequate answer to the 'unity' question. Hence deflationary responses to that question are incorrect. An example of such a response is found in Edward Craig's discussion of Hume: 'Why do I not then take some of your mental states for mine? ... The answer, obviously, is that your perceptions are no part of the input to the device which produces my belief in the unity of my mind. It is not as if there were a mechanism which was allowed first to survey all the perceptions there are, and then parcelled them out.'¹⁷ Of course there is no such 'mechanism'; but 'the device which produces my belief in the unity of my mind' cannot be defined independently of that belief. This is so even if the 'device' is a brain and nervous system rather than a body, since what counts as 'input' to this device depends on the self-ascriptions of the subject. Obviously I do not have IEM-based knowledge of my brain-states, but whatever is the appropriate underlying causal basis for my IEM-based self-ascriptions counts as *my* brain. Subjects are not individuated by brains and nervous systems. So the 'unity question' – 'What is it that unites distinct experiences into experiences had by one person?' – and the 'identity question' – 'What are the criteria for personal identity?' – are inter-dependent.

Now that we have 'psychologized' bodily criteria, what about claim (ii), that bodily criteria thus 'psychologized' do not provide a more fundamental basis for judgements of personal identity than that of traditionally psychological criteria? I have argued that 'bodily' criteria are no more fundamental to personal identity than 'psychological' criteria; if they were, the circularity objection would have some purchase. Moreover, the claim that *q*-memory is unintelligible is in part the claim that we can make no sense of a conflict in particular cases between first- and third-person criteria for personal identity. (It also says that we can make no sense of a conflict within first-person criteria.) These claims help to allay anxieties that the account I am offering, despite protestations to the contrary, involves *au fond* a resurrection of the Cartesian thinking substance. That conclusion seems to beckon if one rejects both Parfitian reductionism and Evans' view that the use of 'I' is a mere artefact of one's way of describing the situation. But I have denied that there is infallible knowledge of one's own identity, and I wish merely to make the negative claim that, given the hypothesis of a conflict between first- and third-person criteria, it would not follow that the latter should be accorded priority. Maybe the correspondence between these criteria is a 'deep contingent fact' of a Wittgensteinian sort; perhaps it

¹⁷ E. Craig, *The Mind of God and the Works of Man* (Oxford UP, 1987), p. 112.

is more like a Kantian presupposition of self-consciousness. Rejection of the circularity objection, and endorsement of a 'unity of consciousness' account of personal identity, are not inconsistent with a caution on science-fictional theorizing about cases where our criteria diverge.

The best direction for an anti-materialist account of personal identity to take does not, therefore, involve breaking the circularity of the memory criterion by postulating *q*-memory, but rather embracing it to reflect the unity of consciousness. The identity of the subject and of the person is together constituted by the ways of knowing that are IEM. It was suspicion of the guaranteed success in self-identification here involved which led Anscombe and (if indeed he did) Wittgenstein to argue that the self is not an object. Their view, stigmatized by Evans as an 'Idealist' conception of the self, might more properly be expressed by saying: the self is an object whose identity is defined by its own self-conception. Perhaps in this way, *transcendental* idealism makes a comeback against naturalism.¹⁸

Durham University

¹⁸ Anscombe, 'The First Person'; Evans p. 212. During the long gestation of this article – from a monograph on self-consciousness, in preparation – I have been grateful for comments from Crispin Wright, John McDowell, Alan Millar, John Skorupski, Lucy O'Brien and E.J. Lowe, from colleagues at a Keele Research Seminar and from the audience at the Cambridge Moral Sciences Club; also from anonymous referees of *The Philosophical Quarterly*.