



**UNIVERSITY OF
ILLINOIS PRESS**

Persons and Their Pasts

Author(s): Sydney Shoemaker

Source: *American Philosophical Quarterly*, Oct., 1970, Vol. 7, No. 4 (Oct., 1970), pp. 269-285

Published by: University of Illinois Press on behalf of the North American Philosophical Publications

Stable URL: <https://www.jstor.org/stable/20009360>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

and University of Illinois Press are collaborating with JSTOR to digitize, preserve and extend access to *American Philosophical Quarterly*

I. PERSONS AND THEIR PASTS

SYDNEY SHOEMAKER

PERSONS have, in memory, a special access to facts about their own past histories and their own identities, a kind of access they do not have to the histories and identities of other persons and other things. John Locke thought this special access important enough to warrant a special mention in his definition of "person," viz., "a thinking, intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places. . . ."¹ In this paper I shall attempt to explain the nature and status of this special access and to defend Locke's view of its conceptual importance. I shall also attempt to correct what now seem to me to be errors and oversights in my own previous writings on this topic.

I

As a first approximation, the claim that persons have in memory a special access to their own past histories can be expressed in two related claims, both of which will be considerably qualified in the course of this paper. The first is that it is a necessary condition of its being true that a person remembers a given past event that he, that same person, should have observed or experienced the event, or known of it in some other direct way, at the time of its occurrence. I shall refer to this as the "previous awareness condition" for remembering.²

The second claim is that an important class of first person memory claims are in a certain respect

immune to what I shall call "error through misidentification." Consider a case in which I say, on the basis of my memory of a past incident, "I shouted that Johnson should be impeached," and compare this with a case in which I say, again on the basis of my memory of a past incident, "John shouted that Johnson should be impeached." In the latter case it could turn out that I do remember someone who looked and sounded just like John shouting that Johnson should be impeached, but that the man who shouted this was nevertheless not John—it may be that I misidentified the person as John at the time I observed the incident, and that I have preserved this misidentification in memory, or it may be that I subsequently misidentified him as John on the basis of what I (correctly) remembered about him. Here my statement would be false, but its falsity would not be due to a mistake or fault of my memory; my memory could be as accurate and complete as any memory could be without precluding this sort of error. But this sort of misidentification is not possible in the former case. My memory report could of course be mistaken, for one can misremember such incidents, but it could not be the case that I have a full and accurate memory of the past incident but am mistaken in thinking that the person I remember shouting was myself. I shall speak of such memory judgments as being immune to error through misidentification with respect to the first

¹ Locke, *Essay Concerning Human Understanding*, Bk. II, Chap. 27, sec. 9 (London, 1912). Italics added.

² In their paper "Remembering" (*The Philosophical Review*, vol. 75 [April, 1966]) C. B. Martin and Max Deutscher express what I call the previous awareness condition by saying that "a person can be said to remember something happening or, in general, remember something directly, only if he has observed or experienced it." Their notion of direct remembering seems to be much the same as Norman Malcolm's notion of "personal memory" (see his "Three Lectures on Memory" in *Knowledge and Certainty* [Englewood Cliffs, N.J., 1963], pp. 203–221). To remember that Caesar invaded Britain I need not have had any experience of the invasion, but no one who lacked such experience could directly or personally remember that Caesar invaded Britain. In this paper I am primarily concerned with memories that are of events, i.e., of something happening, and do not explicitly consider what Malcolm calls "factual memory," i.e., memories that such and such was (or is, or will be) the case, but what I say can be extended to cover all cases of direct or personal memory. Martin and Deutscher hold, and I agree, that remembering something happening is always direct remembering.

There are apparent counterexamples to the previous witnessing condition as I have formulated it. I can be said to remember Kennedy's assassination, which is presumably an event, yet I did not witness or observe it, and the knowledge I had of it at the time was indirect. But while I can be said to remember the assassination, I could hardly be said to remember Kennedy being shot (what I do remember is hearing about it, and the impact this made on me and those around me). Perhaps I can be said to remember the assassination because we sometimes mean by "the assassination" not only the events in Dallas but their immediate effects throughout the nation and world. In any case, when I speak of memories of events in this paper I mean what Martin and Deutscher speak of as memories of something happening.

person pronouns, or other "self-referring" expressions, contained in them.³

I do not contend that all memory claims are immune to error through misidentification with respect to the first person pronouns contained in them. If I say "I blushed when Jones made that remark" because I remember seeing in a mirror someone, whom I took (or now take) to be myself, blushing, it could turn out that my statement is false, not because my memory is in any way incomplete or inaccurate, but because the person I saw in the mirror was my identical twin or double.⁴ In general, if at some past time I could have known of someone that he was ϕ , and could at the same time have been mistaken in taking that person to be myself, then the subsequent memory claims I make about that past occasion will be subject to error through misidentification with respect to the first person pronouns. But if, as is frequently the case, I could not have been mistaken in this way in the past in asserting what I then knew by saying "I *am* ϕ ," then my subsequent memory claim "I *was* ϕ " will be immune to error through misidentification relative to 'I'; that is, it is impossible in such cases that I should accurately remember someone being ϕ but mistakenly take that person to be myself. We might express this by saying that where the present tense version of a judgment is immune to error through misidentification relative to the first person pronouns contained in it, this immunity is *preserved* in memory.⁵ Thus if I claim on the strength of memory that I saw John yesterday, and have a full and accurate memory of the incident, it cannot be the case that I remember someone seeing John but have misidentified that person as myself; my memory claim "I saw John" is subject to error

through misidentification with respect to the term "John" (for it could have been John's twin or double that I saw), but not with respect to 'I'.

II

In his early paper, "Personal Identity," H. P. Grice held that the proposition "One can only remember one's own past experiences" is analytic, but pointed out that this would be analytic in only a trivial way "if 'memory' were to be defined in terms of 'having knowledge of one's own past experiences'." He says that "even if we were to define 'memory' in this sort of way, we should still be left with a question about the proposition, 'one can only have knowledge of one's own past experiences,' which seems to me a necessary proposition."⁶ Now I doubt very much if Grice, or any other philosopher, would now want to hold that is necessarily true, or that it is true at all, that one's own past experiences are the only past experiences of which one can have knowledge. But one does not have to hold this to hold, with Grice, that it is not just a trivial analytic truth that one's own experiences are the only ones that one can remember, i.e., that it is not the case that the necessity of this truth derives merely from the fact that we refuse to *call* someone's having knowledge of a past experience a case of his remembering it unless the past experience belonged to the rememberer himself.

Grice's remarks are explicitly about memory of past experiences, but they raise an important question about all sorts of "event memory." Supposing it to be a necessary truth that the previous witnessing condition must be satisfied in any

³ Although self-reference is typically done with first person pronouns, it can be done with names, and even with definite descriptions—as when De Gaulle says "De Gaulle intends . . .," and the chairman of a meeting says "The Chair recognizes. . . ." In such cases these expressions are "self-referring," not merely because their reference is in fact to the speaker, but also because the speaker intends in using them to refer to himself.

⁴ There is a subtle distinction between this sort of case and cases like the following, which I would not count as a case of error through misidentification. Suppose that Jones says "You are a fool," and I mistakenly think that he is speaking to me. Subsequently I say "I remember Jones calling me a fool," and my statement is false through no fault of my memory. While this is a case of knowing *that* Jones called someone (someone or other) a fool and mistakenly thinking that he was calling me a fool, it is not a case of knowing *of* some particular person that Jones called him a fool but mistakenly identifying that person as oneself. Whereas in the other case we can say, not merely that I know that someone or other blushed, mistakenly think that it was I, but I know *of* some particular person (namely the man I saw in the mirror) that he blushed and have mistakenly identified him as myself.

⁵ I have discussed the immunity to error through misidentification of first person present tense statements in my paper "Self-Reference and Self-Awareness," *The Journal of Philosophy*, vol. 65, 19 (1968). In that paper I made the mistake of associating this feature with the peculiarities of the first-person pronouns. But in fact present tense statements having the appropriate sorts of predicates are immune to error to misidentification with respect to any expressions that are "self-referring" in the sense of footnote 3, including names and definite descriptions. If someone says "De Gaulle intends to remove France from NATO," and is using "De Gaulle" to refer to himself, his statement is in the relevant sense immune to error through misidentification, regardless of whether he is right in thinking his name is "De Gaulle" and that he is the President of France.

⁶ H. P. Grice, "Personal Identity," *Mind*, vol. 50 (1941), p. 344.

genuine case of remembering, is this necessarily true because we would refuse to *count* knowing about a past event as remembering it if the previous awareness condition were not satisfied, or is it necessary for some deeper reason? I think that many philosophers would hold that if this is a necessary truth at all, it is so only in the former way, i.e., in such a way as to make its necessity trivial and uninteresting. Thus G. C. Nerlich, in a footnote to his paper "On Evidence for Identity," says that it is true only of *our* world, not of all possible worlds, that only by being identical with a witness to past events can one have the sort of knowledge of them one has in memory.⁷ On this view it is logically possible that we should have knowledge of past events which we did not ourselves witness, of experiences we did not ourselves have, and of actions we did not ourselves perform, that is in all important respects like the knowledge we have of past events, experiences, and actions in remembering them. If one takes this view it will seem a matter of small importance, if indeed it is true, that the having of such knowledge could not be called "remembering."

It is of course not absolutely clear just what it means to speak of knowledge as being "in all important respects like" memory knowledge, if this is not intended to imply that the knowledge *is* memory knowledge. Presumably, knowledge of past events that is "just like" memory knowledge must not be inferred from present data (diaries, photographs, rock strata, etc.) on the basis of empirical laws and generalizations. But while this is necessary, it is not sufficient. When a person remembers a past event there is a correspondence between his present cognitive state and some past cognitive and sensory state of his that existed at the time of the remembered event and consisted in his experiencing the event or otherwise being aware of

its occurrence.⁸ I shall say that remembering a past event involves there being a correspondence between the rememberer's present cognitive state and a past cognitive and sensory state that was "of" the event.⁹ In actual memory this past cognitive and sensory state is always a past state of the rememberer himself. What we need to consider is whether there could be a kind of knowledge of past events such that someone's having this sort of knowledge of an event does involve there being a correspondence between his present cognitive state and a past cognitive and sensory state that was of the event, but such that this correspondence, although otherwise just like that which exists in memory, does not necessarily involve that past state's having been a state of the very same person who subsequently has the knowledge. Let us speak of such knowledge, supposing for the moment that it is possible, as "quasi-memory knowledge," and let us say that a person who has this sort of knowledge of a past event "quasi-remembers" that past event. Quasi-remembering, as I shall use the term, includes remembering as a special case. One way of characterizing the difference between quasi-remembering and remembering is by saying that the former is subject to a weaker previous awareness condition than the latter. Whereas someone's claim to remember a past event implies that he himself was aware of the event at the time of its occurrence, the claim to quasi-remember a past event implies only that someone or other was aware of it. Except when I indicate otherwise, I shall use the expression "previous awareness condition" to refer to the stronger of these conditions.

Our faculty of memory constitutes our most direct access to the past, and this means, given the previous awareness condition, that our most direct access to the past is in the first instance an access to *our own* past histories. One of the main questions I

⁷ G. C. Nerlich, "On Evidence for Identity," *Australian Journal of Philosophy*, vol. 37 (1959), p. 208.

⁸ I am not here endorsing the view, which I in fact reject, that remembering consists in the having of an image, or some other sort of mental "representation," in which the memory content is in some way encoded. It is sufficient for the existence at *t* of the "cognitive state" of remembering such and such that it be true of the person at *t* that he remembers such and such; I am not here committing myself to any account of what, if anything, someone's remembering such and such "consists in."

⁹ I should make it clear that I am not saying that what we remember is always, or even normally, a past cognitive and sensory state. I am not propounding the view, which is sometimes held but which is clearly false, that "strictly speaking" one can remember only one's own past experiences. I am saying only that if a person remembers an event that occurred at time *t* then at *t* there must have been a corresponding cognitive and sensory state—which the person may or may not remember—that was of that event. It would not be easy to specify just what sort of correspondence is required here, and I shall not attempt to do so. But I take it as obvious that the claim to remember firing a gun requires, for its truth, a different sort of past cognitive and sensory state than the claim to remember hearing someone else fire a gun, and that the latter, in turn, requires a different sort of past cognitive and sensory state than the claim to remember seeing someone fire a gun. Sometimes one remembers a past event but no longer remembers just how one knew of it at the time of its occurrence; in such a case one's memory, because of vagueness and incompleteness, corresponds to a wider range of possible cognitive and sensory states than (say) a memory of seeing the event or a memory of being told about it.

shall be considering in this paper is whether it is conceivable that our most direct access to the past should be a faculty of quasi-remembering which is not a faculty of remembering. Is it conceivable that we should have, as a matter of course, knowledge that is related to past experiences and actions other than our own in just the way in which, as things are, our memory knowledge is related to our own past experiences and actions? In our world all quasi-remembering is remembering; what we must consider is whether the world could be such that most quasi-remembering is not remembering.

Before going on to consider this question I should mention two reasons why I think it important. The first is its obvious bearing on the question of the relationship between the concepts of memory and personal identity. If there can be quasi-remembering that is not remembering, and if remembering can be defined as quasi-remembering that is of events the quasi-rememberer was aware of at the time of their occurrence (thus making it a trivial analytic truth that one can remember an event only if one was previously aware of it), then it would seem that any attempt to define or analyze the notion of personal identity in terms of the notion of remembering will be viciously circular. I shall have more to say about this in Sect. V. But this question also has an important bearing on the question of how a person's memory claims concerning his own past are grounded. In previous writings I have claimed, and made a great deal of the claim, that our memory knowledge of our own past histories, unlike our knowledge of the past histories of other things, is not grounded on criteria of identity.¹⁰ Strawson makes a similar claim in *The Bounds of Sense*, saying that "When a man (a subject of experience) ascribes a current or directly remembered state of consciousness to himself, no use whatever of any criteria of personal identity is required to justify his use of the pronoun 'I' to refer to the subject of that experience." He remarks that "it is because Kant recognized this truth that his treatment of the subject is so greatly superior to Hume's."¹¹ Now it can easily seem that this claim

follows immediately from the fact that remembering necessarily involves the satisfaction of the previous awareness condition. If one remembers a past experience then it has to have been one's own, and from this it may seem to follow that it makes no sense to inquire concerning a remembered experience whether it was one's own and then to try to answer this question on the basis of empirical criteria of identity. But suppose that it were only a trivial analytic truth that remembering involves the satisfaction of the previous awareness condition, and suppose that it were possible to quasi-remember experiences other than one's own. If this were so one might remember a past experience but not know whether one was remembering it or only quasi-remembering it. Here, it seems, it would be perfectly appropriate to employ a criterion of identity to determine whether the quasi-remembered experience was one's own, i.e., whether one remembered it as opposed to merely quasi-remembering it. Thus the question of whether the knowledge of our own identities provided us by memory is essentially non-critical turns on the question of whether it is possible to quasi-remember past actions and experiences without remembering them.

III

There is an important respect in which my characterization of quasi-remembering leaves that notion inadequately specified. Until now I have been ignoring the fact that a claim to remember a past event implies, not merely that the rememberer experienced such an event, but that his present memory is in some way *due to*, that it came about *because of*, a cognitive and sensory state the rememberer had at the time he experienced the event. I am going to assume, although this is controversial, that it is part of the previous awareness condition for memory that a veridical memory must not only correspond to, but must also stand in an appropriate *causal* relationship to, a past cognitive and sensory state of the rememberer.¹² It may seem that if quasi-memory is to

¹⁰ See my book *Self-Knowledge and Self-Identity* (Ithaca, N.Y., 1963), especially Chap. Four, and my paper "Personal Identity and Memory," *Journal of Philosophy*, vol. 56 (1959), pp. 868-882.

¹¹ P. F. Strawson, *The Bounds of Sense* (London, 1966), p. 165.

¹² I owe to Norman Malcolm the point that to be memory knowledge one's knowledge must be in some way due to, must exist because of, a past cognitive and sensory state of oneself—see his "Three Lectures on Memory" (*op. cit.*). Malcolm holds that "due to" does not here express a causal relationship, but I have been persuaded otherwise by Martin's and Deutscher's "Remembering" (*op. cit.*). See also my paper "On Knowing Who One Is" (*Common Factor*, No. 4, 1966), and David Wiggins' *Identity and Spatio-Temporal Continuity* (Oxford, 1967), especially p. 50 ff. The view that there is a causal element in the concept of memory is attacked by Roger Squires in his recent paper "Memory Unchained" (*The Philosophical Review*, vol. 78 [1969] pp. 178-196); I make a very limited reply to this in Sect. V of this paper.

be as much like memory as possible, we should build a similar requirement into the previous awareness condition for quasi-memory, i.e., that we should require that a veridical quasi-memory must not only correspond to, but must also stand in an appropriate causal relationship to, a past cognitive and sensory state of someone or other. On the other hand, it is not immediately obvious that building such a requirement into the previous awareness condition for quasi-memory would not make it equivalent to the previous awareness condition for memory, and thus destroy the intended difference between memory and quasi-memory. But there is no need for us to choose between a previous awareness condition that includes the causal requirement and one that does not, for it is possible and useful to consider both. In the present section I shall assume that the previous awareness condition for quasi-memory does not include the causal requirement, and that it includes nothing more than the requirement that a quasi-memory must, to be a veridical quasi-memory of a given event, correspond in content to a past cognitive and sensory state that was of that event. In the sections that follow I shall consider the consequences of strengthening this condition to include the causal requirement.

The first thing we must consider is what becomes of the immunity of first person memory claims to error through misidentification if we imagine the faculty of memory replaced by a faculty of quasi-memory. As things are now, there is a difference between, on the one hand, remembering an action of someone else's—this might consist, for example, in having a memory of seeing someone do the action—and, on the other hand, remembering *doing* an action, which can be equated with remembering *oneself* doing the action. In the case of quasi-remembering the distinction corresponding to this is that between, on the one hand, the sort of quasi-memory of a past action whose corresponding past cognitive and sensory state belonged to someone who was watching someone else do the action, and, on the other hand, the sort of quasi-memory of a past action whose corresponding past cognitive and sensory state belonged to the very person who did the action. Let us call these, respectively, quasi-memories of an action "from the outside" and quasi-memories of an action "from the inside." Now whereas I can remember an action from the inside only if it was my action, a world in which there is quasi-remembering that is not remembering will be one in which it is not true that any

action one quasi-remembers from the inside is thereby an action he himself did. So—assuming that ours may be such a world—if I quasi-remember an action from the inside, and say on this basis that I did the action, my statement will be subject to error through misidentification; it may be that my quasi-memory of the action is as accurate and complete as it could be, but that I am mistaken in thinking that I am the person who did it. There is another way in which a first person quasi-memory claim could be mistaken through misidentification. If there can be quasi-remembering that is not remembering, it will be possible for a person to quasi-remember an action of his own from the outside. That is, one might quasi-remember an action of one's own as it appeared to someone else who observed it; one might, as it were, quasi-remember it through the eyes of another person. But of course, if I were to quasi-remember someone who looks like me doing a certain action, and were to say on that basis that I did the action, I might be mistaken through no fault of my quasi-memory; it might be that the person who did the action was my identical twin or someone disguised to look like me.

What I have just said about the quasi-remembering of past actions also applies to the quasi-remembering of past experiences and of other mental phenomena. If I remember a past pain from the inside—i.e., remember the pain itself, or remember having the pain, as opposed to remembering seeing someone manifest pain behavior—then the pain must have been mine. But the fact that I *quasi*-remember a pain from the inside will be no guarantee that the pain was mine. Any quasi-memory claim to have been in pain on some past occasion, or to have had a certain thought, or to have made a certain decision, will be subject to error through misidentification.

What is shown by the foregoing is that the immunity of first person memory claims to error through misidentification exists only because remembering requires the satisfaction of the previous awareness condition, and that this feature disappears once we imagine this requirement dropped. Quasi-memory, unlike memory, does not preserve immunity to error through misidentification relative to the first person pronouns. To consider the further consequences of replacing memory with quasi-memory, I must first say something more about memory.

To refer to an event of a certain sort as one that one remembers does not always uniquely identify

it, since one may remember more than one event of a given sort, but it does go some way toward identifying it. In referring to an event in this way one to a certain extent locates it in space and time, even if the description of the event contains no place-names, no names of objects by reference to which places can be identified, and no dates or other temporal indicators. For in saying that one remembers the event one locates it within a spatiotemporal region which is defined by one's own personal history. The spatiotemporal region which is "rememberable" by a given person can be charted by specifying the intervals of past time during which the person was conscious and by specifying the person's spatial location, and indicating what portions of his environment he was in a position to witness, at each moment during these intervals. If someone reports that he remembers an event of a certain kind, we know that unless his memory is mistaken an event of that kind occurred within the spatiotemporal region rememberable by him, and in principle we can chart this region by tracing his history back to its beginning.

Ordinarily, of course, we have far more knowledge than this of the spatiotemporal location of a remembered event, for usually a memory report will fix this position by means of dates, place-names, and other spatial and temporal indicators. But it must be noted that memory claims are subject to error through misidentification with respect to spatial indicators. If a man says "I remember an explosion occurring right in front of that building," it is possible for this to be false even if the memory it expresses is accurate and detailed; the remembered explosion may have occurred, not in front of the building indicated, but in front of another building exactly like it. This remains true no matter how

elaborate and detailed we imagine the memory claim to be. For any set of objects that has actually existed in the world, even if this be as extensive as the set of buildings, streets, parks, bridges, etc., that presently make up New York City, it is logically possible that there should somewhere exist, or that there should somewhere and at some time have existed, a numerically different but exactly similar set of objects arranged in exactly the same way. So memory claims are, in principle, subject to error through misidentification even with respect to such place names as "New York City." Here I am appealing to what Strawson has referred to as the possibility of "massive reduplication."¹³

When a memory report attempts to fix the location of a remembered event by reference to some landmark, we are ordinarily justified in not regarding it as a real possibility that the claim involves error through misidentification owing to the reduplication of that landmark. Certainly we are so justified if the landmark is New York City. But it is important to see why this is so. It is not that we have established that nowhere and at no time has there existed another city exactly like New York; as a self-consistent, unrestricted, negative existential claim, this is something that it would be impossible in principle for us to establish.¹⁴ What we can and do know is that New York is not reduplicated within any spatiotemporal region of which anyone with whom we converse can have had experience. Whether or not New York is reduplicated in some remote galaxy or at some remote time in the past, we know that the man who claims to remember doing or experiencing something in a New York-like city cannot have been in any such duplicate. And from this we can conclude that if he does remember doing or ex-

¹³ P. F. Strawson, *Individuals* (London 1959), p. 20.

¹⁴ It will perhaps be objected that the dictum that unrestricted negative existential claims are unfalsifiable in principle is brought into question by the possibility that we might discover—what some cosmologists hold there is good reason for believing—that space and past time are finite. If we discovered this, why shouldn't we be able, at least in principle, to establish that at no place does there exist, and at no time in the past has there existed, a duplicate of New York?

One way of countering this objection would be to introduce the possibility, which has been argued by Anthony Quinton in his paper "Spaces and Times" (*Philosophy*, vol. 57 [1962] pp. 130–141), of there being a multiplicity of different and spatially unrelated spaces. Establishing that there is no duplicate of New York in our space would not establish that there is no space in which there is such a duplicate, and if it is possible for there to be multiplicity of spaces there would seem to be no way in which the latter could be established.

But we needn't have recourse to such recondite possibilities in order to counter the objection, if it is viewed as an objection to my claim that it is the fact that remembering involves the satisfaction of the previous awareness condition that makes it possible for us to rule out the possibility that memory claims are false through misidentification owing to the reduplication of landmarks. For to discover that space or past time is finite, and that massive reduplication does not occur, one would have to have a vast amount of empirical information about the world, including information about the histories of particular things. But, as I think the remainder of my discussion should make clear, one could not be provided with such information by memory (or by quasi-memory) unless one were *already* entitled in a large number of cases to refer to particular places and things in one's memory reports without having to regard it as possible that one's references were mistaken owing to massive reduplication. So this entitlement would have to precede the discovery that space and past time are finite, and could not depend on it.

periencing something in a New York-like city, then it was indeed in New York, and not in any duplicate of it, that the remembered action or event occurred. But we can conclude this only because remembering involves the satisfaction of the previous awareness condition.

Even when a landmark referred to in someone's memory claim is reduplicated within the spatiotemporal region rememberable by that person, we can often be confident that the claim does not involve error through misidentification. Suppose that someone locates a remembered event, say an explosion, by saying that it occurred in front of his house, and we know that there are many houses, some of which he has seen, that are exactly like his. If he reported that he had simply found himself in front of his house, with no recollection of how he had gotten there, and that after seeing the explosion he had passed out and awakened later in a hospital, we would think it quite possible that he had misidentified the place at which the remembered explosion occurred. But suppose instead that he reports that he remembers walking home from work, seeing the explosion in front of his house, and then going inside and being greeted by his family. Here a misidentification of the place of the explosion would require the reduplication, not merely of his house, but also of his family, his place of work, and the route he follows in walking home from work. We could know that no such reduplication exists within the spatiotemporal region of which he has had experience, and could conclude that his report did not involve an error through misidentification. But again, what would enable us to conclude this is the fact that remembering involves the satisfaction of the previous awareness condition.

Presumably, what justifies any of us in using such expressions as "New York" and "my house" in his own memory reports are considerations of the same kind as those that justify others in ruling out the possibility that claims containing such expressions involve error through misidentification. What justifies one is the knowledge that certain sorts of reduplication do not in fact occur within the spatiotemporal regions of which any of us have had experience. Normally no such justification is needed for the use of 'I' in memory reports; this is what is involved in saying that memory claims are normally immune to error through misidentification relative to the first person pronouns. But what makes such a justification possible in the case of "New York" is the same as what makes it un-

necessary in the case of 'I', namely the fact that remembering involves the satisfaction of the previous awareness condition. So it is because of this fact that remembering can provide us, not merely with the information that an event of a certain sort has occurred somewhere or other in the vicinity of persons and things satisfying certain general descriptions, but with the information that such an event occurred in a certain specified place, in a certain specifiable spatial relationship to events presently observed, and in the vicinity of certain specified persons or things. But this is also to say that it is this fact about remembering that makes it possible for us to know that an object or person to which one remembers something happening is, or is not, identical with an object or person presently observed. And it will emerge later that it is also this fact about remembering that makes it possible to know that different memories are, or are not, of events in the history of a single object or person.

But now let us consider the consequences of replacing the faculty of memory by a faculty of quasi-memory. Quasi-remembering does not necessarily involve the satisfaction of the previous awareness condition, and first person quasi-memory claims are, as we have seen, subject to error through misidentification. It is a consequence of this that even if we are given that someone's faculty of quasi-memory is highly reliable, in the sense that when he seems to quasi-remember an event of a certain sort he almost always does quasi-remember such an event, nevertheless his quasi-memory will provide neither him nor us with any positive information concerning the spatial location of the events he quasi-remembers, or with any information concerning the identity, or concerning the history, of any object or person to which he quasi-remembers something happening. The fact that he quasi-remembers an event of a certain sort will not provide us with the information that such an event has occurred within the spatiotemporal region of which he has had experience. But in consequence of this, if he attempts to locate the quasi-remembered event by reference to some object or place known to us, e.g., New York or Mt. Everest, it is impossible for us to rule out on empirical grounds the possibility that his claim involves error through misidentification owing to the reduplication of that object or place. To rule this out we would have to have adequate grounds for asserting, not merely that there is no duplicate of New York (say) in the spatiotemporal region of

which he has had experience, but that at no place and time has there been a duplicate of New York. And this we could not have.¹⁵ But this means that in expressing his quasi-memories he could not be justified in using such expressions as "New York" and "Mt. Everest," or such expressions as "I," "this," and "here," to refer to the places, persons, and things in or to which he quasi-remembers certain things happening. The most he could be entitled to assert on the basis of his quasi-memories would be a set of general propositions of the form "An event of type ϕ at some time occurred in the history of an object of type A while it stood in relations $R_1, R_2, R_3 \dots$ to objects of types $B, C, D \dots$ " And given only a set of propositions of this sort, no matter how extensive, one could not even begin to reconstruct any part of the history of the world; one could not even have grounds for asserting that an object mentioned in one proposition in the set was one and the same as an object mentioned in another proposition of the set.

So far I have been ignoring the fact that the events and actions we remember generally have temporal duration, and the fact that we sometimes remember connected sequences of events and actions lasting considerable lengths of time. What will correspond to this if remembering is replaced with quasi-remembering? If someone says "I remember doing X and then doing Y ," it would make no sense to say to him, "Granted that your memory is accurate, and that such a sequence of actions did occur, are you sure that it was one and the same person who did both X and Y ?" But now suppose that someone says "I quasi-remember doing X and then doing Y ," and that the world is such that there is quasi-remembering that is not remembering. Here it is compatible with the accuracy of the man's quasi-memory that he should be mistaken in thinking that he himself did X and Y . And as I shall now try to show, it must also be compatible with the accuracy of this man's quasi-memories that he should be mistaken in thinking even that one and the same person did both X and Y .

Suppose that at time t_1 a person, call him A , does action Y and has while doing it a quasi-memory from the inside of the immediately previous occurrence of the doing of action X . A 's having this quasi-memory of the doing of X is of course compatible with X 's having been done by some-

one other than himself. At t_1 A 's cognitive state includes this quasi-memory from the inside of the doing of X together with knowledge from the inside of the doing of Y ; we might say that it includes knowledge from the inside of the action sequence X -followed-by- Y . But now suppose that at a later time t_2 someone, call him B , has a quasi-memory corresponding to the cognitive state of A at t_1 . It would seem that B 's quasi-memory will be a quasi-memory from the inside of the action sequence X -followed-by- Y . This quasi-memory will be veridical in the sense that it corresponds to a past cognitive state that was itself a state of knowledge, yet its being veridical in this way is compatible with X and Y having been done by different persons. If A were mistakenly to assert at t_1 that X and Y were done by the same person, his mistake would not be due to a faulty quasi-memory. And if B 's cognitive state at t_2 corresponds to A 's cognitive state at t_1 , then if B were mistaken at t_2 in thinking that X and Y were done by the same person, this mistake would not be due to a faulty quasi-memory.

If, as I have been arguing, someone's quasi-remembering from the inside the *action* sequence X -followed-by- Y provides no guarantee that X and Y were done by the same person, then by the same reasoning someone's quasi-remembering the *event* sequence X -followed-by- Y provides no guarantee that X and Y were witnessed by the same person, and therefore no guarantee that they occurred in spatial proximity to one another. But any temporally extended event can be thought of as a succession of temporally and spatially contiguous events; e.g., a stone's rolling down a hill can be thought of as consisting in its rolling half of the way down followed by its rolling the other half of the way. Suppose, then, that someone has a quasi-memory of the following event sequence: stone rolling from top of hill to middle followed by stone rolling from middle of hill to bottom. If we knew this to be a memory, and not just a quasi-memory, we would know that if it is veridical then one and the same person observed both of these events, one immediately after the other, and this together with the contents of the memory could guarantee that one and the same hill and one and the same stone were involved in both, and that a single stone had indeed rolled all the way down a hill. But the veridicality of this quasi-memory *qua* quasi-memory would be compatible with these events

¹⁵ The point made in the preceding footnote can now be expressed by saying that even if we, who have the faculty of memory, could establish that at no place and time has there been a duplicate of New York, this could not be established by someone whose faculty of knowing the past was a faculty of quasi-memory.

having been observed by different persons, and with their involving different stones and different hills; it would be compatible with no stone's having rolled all of the way down any hill. And since any temporally extended event can be thought of as a succession of temporally and spatially contiguous events, it follows that someone's quasi-remembering what is ostensibly a temporally extended event of a certain kind is always compatible with there actually being no such event that he quasi-remembers, for it is compatible with his quasi-memory being, as it were, compounded out of quasi-memories of a number of different events that were causally unrelated and spatiotemporally remote from one another. The knowledge of the past provided by such a faculty of quasi-memory would be minimal indeed.¹⁶

IV

But now we must consider the consequences of strengthening the previous awareness condition for quasi-remembering to include the requirement that a veridical quasi-memory must not only correspond to, but must also stand in an appropriate causal relationship to, a past cognitive and sensory state of someone or other. Clearly, much of what I have said about quasi-remembering ceases to hold once its previous awareness condition is strengthened in this way. If, as is commonly supposed, causal chains must be spatiotemporally continuous, then if quasi-memory claims implied the satisfaction of this strengthened previous

awareness condition they would, when true, provide some information concerning the location of the quasi-remembered events and actions. We would know at a minimum that the spatiotemporal relationship between the quasi-remembered event and the making of the quasi-memory claim is such that it is possible for them to be linked by a spatiotemporally continuous causal chain, and if we could trace the causal ancestry of the quasi-memory we could determine precisely when and where the quasi-remembered event occurred. Thus if we construe the previous awareness condition of quasi-memory as including this causal requirement, it seems that a faculty of quasi-remembering could enable us to identify past events and to reidentify persons and things, and it seems at first glance (though not, I think, on closer examination) that it would enable us to do this without giving us a special access to our own past histories.

It must be stressed that this strengthened previous awareness condition is an improvement on the weaker one *only* on the assumption that causal chains (or at any rate the causal chains that link cognitive and sensory states with subsequent quasi-memories) must be spatiotemporally continuous, or at least must satisfy a condition similar to spatiotemporal continuity. If the sort of causality operating here allowed for action at a spatial or temporal distance, and if there were no limit on the size of the spatial or temporal gaps that could exist in a causal chain linking a cognitive and sensory state with a subsequent quasi-memory, then the claim that a quasi-memory originated

¹⁶ It may be objected that I have overlooked one way in which a quasi-rememberer might begin to reconstruct his own past history, and the histories of other things, from the information provided him by his quasi-memories. The quasi-rememberer's difficulties would be solved if he had a way of sorting out those of his quasi-memories that are of his own past, i.e., are memories, from those that are not. But it may seem that the quasi-rememberer could easily tell which of his quasi-memories of the very *recent* past are of his own past, namely by noting which of them have contents very similar to the contents of his *present* experiences; e.g., if he quasi-remembers from the inside the very recent seeing of a scene that resembles very closely the scene he presently sees, it may seem that he can justifiably conclude that the quasi-remembered seeing was his own. And it may seem that by starting in this way he could trace back his own history by finding among his quasi-memories a subset of situations that form a spatiotemporally continuous series of situations, that series terminating in the situation he presently perceives.

This objection assumes that the quasi-rememberer can know the degree of recentness of the situations of which he has quasi-memories, but I shall not here question this assumption. What I shall question is the assumption that if the quasi-rememberer knows that a quasi-remembered scene occurred only a moment or so ago, and that it closely resembles the scene he presently sees, he is entitled to believe that it is numerically the same scene as the one he presently sees and that in all probability it was he who saw it. For of course it could be the case that there is somewhere else a duplicate of the scene he sees, and that his quasi-memory is of that duplicate. It will perhaps be objected that while this is logically possible (given the possibility of quasi-remembering that is not remembering), it is highly improbable. But while it may be intrinsically improbable that a highly complicated situation should be reduplicated within some limited spatiotemporal area, it does not seem intrinsically improbable that such a situation should be reduplicated somewhere or other in the universe—unless the universe is finite, which is something the quasi-rememberer could have no reason for believing (see footnotes 14 and 15). Moreover, one could not be in a position to know how rare or frequent such reduplication is in fact, and therefore how likely or unlikely it is that a given situation is reduplicated, unless one already had a way of reidentifying places and things. So the quasi-rememberer could not be in a position to know this, for he could have a way of reidentifying places and things only if he were already in a position to rule out reduplication as improbable.

in a corresponding cognitive and sensory state would be as unfalsifiable, and as uninformative, as the claim that it corresponds to a past cognitive and sensory state of someone or other.

To consider the consequences of strengthening the previous awareness condition for quasi-memory in the way just suggested I shall have to introduce a few technical expressions. First, I shall use the expressions "quasi_c-remember" and "quasi_c-memory" when speaking of the sort of quasi-remembering whose previous awareness condition includes the causal requirement. Second, I shall use the term "M-type causal chain" to refer to the sort of causal chain that must link a quasi_c-memory with a corresponding past cognitive and sensory state if they are to be "of" the same event, or if the former is to be "of" the latter. Since quasi_c-remembering is to be as much like remembering as is compatible with the failure of the strong previous awareness condition, M-type causal chains should resemble as much as possible the causal chains that are responsible for actual remembering, i.e., should resemble them as much as is compatible with their sometimes linking mental states belonging to different persons. At any given time a person can be said to have a total mental state which includes his memories or quasi_c-memories and whatever other mental states the person has at that time. Let us say that two total mental states, existing at different times, are directly M-connected if the later of them contains a quasi_c-memory which is linked by an M-type causal chain to a corresponding cognitive and sensory state contained in the earlier. And let us say, by way of giving a recursive definition, that

two total mental states are M-connected if either (1) they are directly M-connected, or (2) there is some third total mental state to which each of them is M-connected.¹⁷

Now there are two cases we must consider. Either the world will be such, or it will not, that a total mental state existing at a particular time can be M-connected with at most one total mental state existing at each other moment in time. Or, what comes to the same thing, either the world will be such, or it will not, that no two total mental states existing at the same time can be M-connected. Let us begin by considering the case in which the former of these alternatives holds. This is the case that will exist if there is no "branching" of M-type causal chains, i.e., if it never happens that an M-type causal chain branches into two such chains which then produce quasi_c-memories belonging to different and simultaneously existing total mental states, and if it never happens that different M-type causal chains coalesce and produce in a single total mental state quasi_c-memories whose corresponding past cognitive and sensory states belonged to different and simultaneously existing total mental states. This is presumably the situation that exists in the actual world. And I think that in any world in which this situation exists M-connected total mental states will be, to use a term of Bertrand Russell's, "copersonal," i.e., states of one and the same person, and quasi_c-remembering will reduce to remembering. There seems to me to be at least this much truth in the claim that memory is constitutive of personal identity.¹⁸ (But more about this in Sect. V.)

Now let us consider the case in which M-type

¹⁷ It is worth mentioning that if quasi_c-remembering is to be as much like remembering as possible then not just any causal chain linking a past cognitive and sensory state with a subsequent quasi_c-memory can be allowed to count as an M-type causal chain. For as Martin and Deutscher (*op. cit.*) point out, there are various sorts of cases in which a man's knowledge of a past event is causally due to his previous experience of it but in which the causal connection is obviously not of the right kind to permit us to say that he remembers the event. E.g., I have completely forgotten the event, but know of it now because you told me about it, and you came to know about it through my telling you about it prior to my forgetting it. It is easier to decide in particular cases whether the causal connection is "of the right kind" than it is to give a general account of what it is for the causal connection to be of the right kind, i.e., what it is for there to be an M-type causal chain. I shall not attempt to do the latter here. The notion of an M-type causal chain would of course be completely useless if it were impossible to determine in any particular case whether the causal connection is "of the right kind" without already having determined that the case is one of remembering—but I shall argue in Sect. V that this is not impossible.

¹⁸ In his paper "Bodily Continuity and Personal Identity: A Reply" (*Analysis*, vol. 21 [1960] pp. 42–48), B. A. O. Williams says that "identity is a one-one relation, and . . . no principle can be a criterion of identity for things of type *T* if it relies on what is logically a one-many or many-many relation between things of type *T*," and remarks that the relation "being disposed to make sincere memory claims which exactly fit the life of" is a many-one relation and "hence cannot possibly be adequate in logic to constitute a criterion of identity" (pp. 44–45). Now it may seem that my version of the view that memory is a criterion of personal identity is open to the same objection, for if M-type causal chains can branch and coalesce then the relation "has a quasi-memory which is linked by an M-type causal chain with a cognitive and sensory state of" is not logically a one-one relation. But while this relationship is not logically one-one, the relationship "has a quasi-memory which is linked by a *non-branching* M-type causal chain with a cognitive and sensory state of" is logically one-one, and it is the holding of the latter relationship that I would hold to be a criterion, in the sense of being a sufficient condition, for personal identity.

causal chains do sometimes branch, and in which, as a result, it can happen that two or more simultaneously existing total mental states are M-connected. Here we cannot claim that if two total mental states are M-connected they are thereby copersonal without committing ourselves to the unattractive conclusion that a person can be in two different places, and can have two different total mental states, at one and the same time. But it is still open to us to say that if a total mental state existing at time t_1 and a total mental state existing at time t_2 are M-connected then they are copersonal *unless* the M-type causal chain connecting them branched at some time during the interval t_1 – t_2 . If we can say this, as I think we can, then even in a world in which there is branching of M-type causal chains the fact that a person quasi-remembers a past event or action would create a presumption that he, that same person, experienced the event or did the action, and therefore a presumption that the quasi-memory was actually a memory. This presumption would stand as long as there was no evidence that the M-type causal chain linking the past action or experience with the subsequent quasi-memory had branched during the interval between them.

Worlds of the sort we are now considering, i.e., worlds in which M-type causal chains sometimes branch, could be of several kinds. Consider first a world in which people occasionally undergo fission or fusion; i.e., people sometimes split, like amoebas, both offshoots having quasi-memories of the actions done prior to the fission by the person who underwent it, and two people sometimes coalesce into a single person who then has quasi-memories of both of their past histories. Here we cannot say that a person did whatever actions he quasi-remembers from the inside without running afoul of Leibniz' Law and the principle of the transitivity of identity. But we can say something close to this. Suppose that someone, call him Jones, splits into two persons, one of whom is I and the other is

someone I shall call Jones II. Both Jones I and II have quasi-memories from the inside of Jones's past actions, and no one else does. If anyone now alive is identical with Jones it is either myself or Jones II, and any objection to saying that I am Jones is equally an objection to saying that Jones II is Jones. I think that we can say here that I am identical with Jones if anyone now alive is identical with him. Or suppose that two people, call them Brown and Smith, coalesce, resulting in me. I have quasi-memories from the inside of Brown's actions and also of Smith's actions. There are serious objections to identifying me with either Brown or Smith, but it seems clear here that if anyone now alive is identical with either Brown or Smith, I am. So in such a world the following principle holds: if at time t a person A quasi-remembers a past action X from the inside then A is identical with the person who did X if anyone alive at t is identical with him.¹⁹

But I think that we can imagine a world in which this principle would not hold. In the case in which two persons coalesce the M-type causal chains involved might be represented by a river having two "forks" of equal width. Suppose that instead of this we have an M-type causal chain, or a connected set of such causal chains, that could be represented by a river having several small tributaries. For example, suppose, very fancifully, that memories were stored, by some sort of chemical coding, in the blood rather than in brain cells, and that as a result of being given a blood transfusion one sometimes acquired quasi-memories "from the inside" of a few of the actions of the blood donor. Here the blood transfusion would be a "tributary" into what apart from its tributaries would be the sort of M-type causal chain that occurs in the history of a single person. Now I do not think that we would deny that A , existing at time t_2 was the same person as B , who existed at an earlier time t_1 , merely because A quasi-remembers from the inside, as the result of a blood transfusion, an

¹⁹ A. N. Prior has defended the view that in cases of fission *both* offshoots can be identified with the original person, although not with each other. This of course involves modifying the usual account of the logical features of identity. See his "Opposite Number" (*Review of Metaphysics*, vol. 11 [1957] pp. 196–201), and his "Time, Existence and Identity" (*Proceedings of the Aristotelian Society*, 1965–1966). Roderick Chisholm takes a very different view. Considering the supposition that "you knew that your body, like that of an amoeba, would one day undergo fission and that you would go off, so to speak, in two different directions," he says "it seems to me, first, that there is no possibility whatever that *you* would be *both* the person on the right and the person on the left. It seems to me, secondly, that there *is* a possibility that you would be one or the other of those two persons" ("The Loose and Popular and the Strict and Philosophical Senses of Identity," in *Perception and Personal Identity*, ed. by Norman S. Care and Robert H. Grimm [Cleveland, 1969], p. 106). It is not clear to me whether Chisholm would hold that one (but not both) of the offshoots might be me if the memories of each stood in the same causal relationships to my actions and experiences as the memories of the other, and if each resembled me, in personality, appearance, etc., as much as the other. If so, I would disagree.

action at t_1 that was not done by B . Nor would we deny that another person C , the blood donor, is the person who did that past action merely because there is someone other than himself, namely A , who quasi_c-remembers it from the inside. So here it would not be true that if at time t a person quasi_c-remembers a past action from the inside then he is identical with the person who did it if anyone existing at t is identical with the person who did it.

Yet even in such a world it seems essential that in any total mental state the memories, i.e., the quasi_c-memories produced by the past history of the person whose total mental state it is, should outnumber the quasi_c-memories produced by any given tributary. If the quasi_c-memories produced by a given tributary outnumbered the memories then surely the tributary would not be a tributary at all, but would instead be the main stream. But this implies that if a person quasi_c-remembers an action from the inside then, in the absence of evidence to the contrary, he is entitled to regard it as more likely that the action was done by him than that it was done by any other given person. And this, taken together with my earlier point that if someone quasi_c-remembers an action from the inside there is a presumption that he is the person who did it, gives us a sense in which quasi_c-memory can be said to provide the quasi_c-rememberer with "special access" to his own past history. This is of course a much weaker sense of "special access" than that explained in Sect. I—but in this sense it will be true in *any* possible world, and not merely in ours, that people have a special access to their own past histories.

V

In the preceding sections it was assumed that remembering, as opposed to (mere) quasi_c-remembering, necessarily involves the satisfaction of the strong previous awareness condition; that is, it was assumed that in any genuine case of event memory the memory must correspond to a past cognitive and sensory state of the rememberer himself. And this is commonly supposed in discussions of memory and personal identity. But it is not really clear that this assumption is correct. For consider again the hypothetical case in which a man's body "splits" like an amoeba into two physiologically identical bodies, and in which both offshoots produce memory claims corresponding to

the past life of the original person. Or, to take a case that lies closer to the realm of real possibility, consider the hypothetical case in which a human brain is split, its two hemispheres are transplanted into the newly vacated skulls of different bodies, and both transplant recipients survive, regain consciousness, and begin to make memory claims that correspond to the past history of the brain "donor."²⁰ In neither case can we identify both of the physiological offshoots of a person with the original person, unless we are willing to take the drastic step of giving up Leibniz' Law and the transitivity of identity. But it is clear that it would be wrong to say that each of the offshoots remembers the actions, experiences, etc., of the original person? There is, to be sure, an awkwardness about saying that each offshoot remembers *doing* an action done by the original person, for this seems to imply that an action done by one and only one person was done by each of the two non-identical offshoots. But perhaps we can say that each of the offshoots does remember the action "from the inside." In our world, where such bizarre cases do not occur, the only actions anyone remembers from the inside are those that he himself performed, so it is not surprising that the only idiomatic way of reporting that one remembers an action from the inside is by saying that one remembers doing the action. But this need not prevent us from describing my hypothetical cases by saying that both offshoots do remember the actions of the original person, and it does not seem to me unnatural to describe them in this way. If this is a correct way of describing them, then perhaps my second sort of quasi-remembering, i.e., quasi_c-remembering, turns out to be just remembering, and the previous awareness condition for remembering turns out to be the causal requirement discussed in the preceding section rather than the stronger condition I have been assuming it to be.

If the suggestion just made about the conditions for remembering is correct, the logical connection between remembering and personal identity is looser than I have been supposing it to be. Yet adopting this suggestion does not prevent one from defending the claim that remembering is constitutive of and criterial for personal identity; on the contrary, this makes it possible to defend the letter of this claim, and not just its spirit, against the very common objection that any attempt to analyze

²⁰ See Wiggins, *op. cit.*, p. 53, where such a case is discussed.

personal identity in terms of memory will turn out to be circular.

Bishop Butler objected against Locke's account of personal identity that "one should really think it self-evident, that consciousness of personal identity presupposes, and therefore cannot constitute, personal identity, any more than knowledge, in any other case, can constitute truth, which it presupposes."²¹ More recently several writers have argued that while "*S* remembers doing *A*" entails "*S* did *A*" (and so entails "*S* is identical with the person who did *A*"), this is only because "*S* remembers doing *A*" is elliptical for "*S* remembers himself doing *A*."²² To offer as a partial analysis of the notion of personal identity, and as a criterion of personal identity, the formula "If *S* remembers (himself) doing action *A*, *S* is the same as the person who did *A*" would be like offering as a partial definition of the word "red," and as a criterion of redness, the formula "If *S* knows that *X* is red, then *X* is red." In both cases the concept allegedly being defined is illicitly employed in the formulation of the defining condition. Likewise, it has been argued that while someone's remembering a past event is a sufficient condition of his being identical with a witness to the event, we cannot use the former as a criterion for the latter, since in order to establish that a person really does remember a given past event we have to establish that he, that very person, was a witness to the event. And if this is so, the formula "If *S* remembers *E*, *S* is identical with someone who witnessed *E*" will be circular if offered as a partial analysis of the concept of personal identity.²³

Such objections assume that remembering involves the satisfaction of the strong previous awareness condition, and they can be avoided on the assumption that the previous awareness condition is weaker than this, e.g., is that given for quasi-remembering in Sect. IV. Or, better, they can be avoided if we explicitly use "remember" in a "weak" sense ("remember_w") rather than in a "strong" sense ("remember_s"), the strength of the

sense depending on the strength of the associated previous awareness condition. Although there are perhaps other possibilities, let us take "remember_w" to be synonymous with "quasi-remember." Clearly, to establish that *S* remembers_w event *E* (or remembers_w action *A* from the inside) it is not necessary to establish that *S* himself witnessed *E* (or did *A*), for it will be enough if *S* is the offshoot of someone who witnessed *E* (did *A*). And while we cannot claim that statements about what events or actions a man remembers_w logically entail statements about his identity and past history, this does not prevent the truth of the former from being criterial evidence for, and from being partially constitutive of, the truth of the latter. For we can still assert as a logical truth that if *S* remembers_w event *E* (or remembers_w action *A* from the inside), and if there has been no branching of M-type causal chains during the relevant stretch of *S*'s history, then *S* is one of the witnesses of *E* (is the person who did *A*). Here we avoid the circularity that Butler and others have thought to be involved in any attempt to give an account of personal identity, and of the criteria of personal identity, in terms of memory.

In the actual world, people remember_s whatever they remember_w, and this makes it difficult to settle the question of whether it is the weak or the strong sense of "remember" that is employed in ordinary discourse. It is possible that this question has no answer; since branching of M-type causal chains does not in fact occur, and is seldom envisaged, people have had no practical motive for distinguishing between the strong and the weak senses of "remember." But I do not think that this question is especially important. We can defend the spirit of the claim that memory is a criterion of personal identity without settling this question, although in order to defend the letter of that claim we must maintain that in its ordinary use "remember" means "remember_w."

At this point I should say something about why it is important to insist on the claim that there is a

²¹ Joseph Butler, "Of Personal Identity," First Dissertation to the *Analogy of Religion*. Reprinted in Flew, ed., *Body, Mind and Death*, (New York, 1964), pp. 166–172.

²² See A. J. Ayer, *The Problem of Knowledge*, (Harmondsworth, Middlesex, 1956), p. 196, and B. A. O. Williams, "Personal Identity and Individuation," in Gustavsen (ed.) *Essays in Philosophical Psychology* (New York, 1964), pp. 327–328 (originally published in the *Proceedings of the Aristotelian Society*, vol. 57, 1956–57).

²³ See Williams, *op. cit.*, p. 329, and my *Personal Identity and Memory*, pp. 869–870 and p. 877. In the latter, and in *Self-Knowledge and Self-Identity*, I attempt to reduce the force of this objection by arguing that it is a "conceptual truth" that memory claims are generally true, and that we can therefore be entitled to say that a person remembers a past event without already having established, or having inductive evidence, that some other criterion of personal identity (one not involving memory) is satisfied. This way of handling the objection no longer seems to me satisfactory.

causal element in the notion of memory. For this claim has recently come under attack.²⁴ It has been argued that the notion of memory should be analyzed in terms of the *retention*, rather than the causation, of knowledge, and that the notion of retention is not itself a causal notion. Now I have no objection to saying that remembering_s consists in the retention of knowledge. But I believe that unless we understand the notion of retention, as well as that of memory, as involving a causal component, we cannot account for the role played by the notion of memory, or even the concept of similarity, in judgments of personal identity.

Here it will be useful to consider a hypothetical case I have discussed at some length elsewhere.²⁵ Let us suppose that the brain from the body of one man, Brown, is transplanted into the body of another man, Robinson, and that the resulting creature—I call him “Brownson”—survives and upon regaining consciousness begins making memory claims corresponding to the past history of Brown rather than that of Robinson. We can also suppose that Brownson manifests personality traits strikingly like those previously manifested by Brown and quite unlike those manifested by Robinson. Although Brownson has Robinson’s (former) body, I doubt if anyone would want to say that Brownson is Robinson, and I think that most people would want to say that Brownson is (is the same person as) Brown.

But what can we offer as evidence that Brownson is Brown? Clearly the mere correspondence of Brownson’s ostensible memories to Brown’s past history, and the similarity of Brownson’s personality to Brown’s, is far from being sufficient evidence. And it is equally clear that the notion of the *retention* of knowledge and traits is of no use here. To be sure, once we take ourselves to have established that Brownson is Brown we can say that Brownson retains knowledge, and also personality traits, acquired by Brownson in the past. But the latter assertion presupposes the identity of Brownson and Brown, and cannot without circularity be offered as evidence for it. Indeed, the circularity is the same as what would be involved in offering as evidence of this identity the fact that Brown-

son remembers_s Brown’s past experiences and actions.

We do not, however, beg the question about identity if we take Brownson’s possession of what used to be Brown’s brain, together with the empirical facts about the role played by the brain in memory, as establishing that Brown’s ostensible memories are directly M-connected with Brown’s past actions and experiences, i.e., are causally related to them in essentially the same ways as people’s memories are generally connected with their own past experiences and actions. This in turn establishes that Brownson quasi_c-remembers, and so remembers_w, Brown’s past experiences and actions. And from this in turn, and from the fact that we have good reason to suppose that no other person’s memories are M-connected with Brown’s past history in this way, i.e., that there has been no “branching” of M-type causal chains, we can conclude that Brownson is Brown.²⁶

We can reason in this way only if we can assert that there is a causal connection between Brownson’s past history and Brownson’s ostensible memories. And this, it seems to me, we are clearly entitled to do. Given that Brownson has Brown’s former brain, there is every reason to think that had Brown’s history been different in certain ways, there would (*ceteris paribus*) be corresponding differences in what Brownson ostensibly remembers. I can see no reason for doubting that such counterfactuals assert causal connections. Similar remarks can be made about the similarity between Brownson’s and Brown’s personality traits. Given that Brownson has Brown’s former brain, we have reason to think that had Brown developed a different set of personality traits, Brownson would (*ceteris paribus*) have those personality traits rather than the ones he has. And while we cannot naturally speak of Brown’s having a certain trait at one time as causing Brownson to have the same trait at a subsequent time, we can speak of the former as being an important part of a causally sufficient condition for the latter. It is only where we suppose that the traits of things at different times are causally related in this way that we are entitled to take the similarity of something at one

²⁴ See Squires’ “Memory Unchained,” *op. cit.*

²⁵ *Self-Knowledge and Self-Identity*, pp. 23–25 and 245–47.

²⁶ In *Self-Knowledge and Self-Identity* I held that saying that Brownson is Brown would involve making a “decision” about the relative weights to be assigned to different criteria of personal identity, and that in the absence of such a decision there is no right answer to the question whether Brownson is Brown. I have come to believe that there is a right answer to this question, namely that Brownson is Brown, and that my former view overlooked the importance of the causal component in the notion of memory—see my treatment of this example in “On Knowing Who One Is,” *op. cit.*

time and something at another time as evidence of identity.

VI

We are now in a position to reassess the view, mentioned in Sect. II, that the knowledge of our own pasts and our own identities provided us by memory is essentially "noncriterial." If I remember_s an action or experience from the inside, and know that I do, it makes no sense for me to inquire whether that action or experience was my own. But it seems logically possible that one should remember_w an action or experience from the inside (i.e., quasi_c-remember it) without remembering_s it. So if one remembers_w an action or experience from the inside it can make sense to inquire whether it was one's own (whether one remembers_s it), and it would seem offhand that there is no reason why one should not attempt to answer this question on the basis of criteria of personal identity.

But while an action I remember_w from the inside can fail to be mine, there is only one way in which this can happen, namely through there having been branching in the M-type causal chain linking it with my present memory. So in asking whether the action was mine, the only question I can significantly be asking is whether there was such branching. If I go on to verify that there was no branching, I thereby establish that a sufficient criterion of personal identity is satisfied. If instead I conclude on inductive grounds that there was no branching, relying on my general knowledge that M-type causal chains seldom or never branch (or that it is physiologically impossible for them to do so), I thereby conclude that a sufficient criterion of personal identity is satisfied. But an important part of what the satisfaction of this criterion consists in, namely my remembering_w the past action from the inside, is not something I establish, and not something I conclude on inductive grounds, but is something I necessarily presuppose in inquiring concerning my relation to the remembered_w action. In cases where one remembers_w a past action from the inside, and knows of it only on that basis, one cannot significantly inquire concerning it whether one does remember_w it—for as I tried to bring out in my discussion of quasi-remembering, there is no way of knowing the past that stands to remembering_w as remembering_w stands to remembering_s, i.e., is such that one can know of a past event in this way and regard it

as an open question whether in so knowing of it one is remembering_w it. So in such cases the satisfaction of this part of the memory criterion for personal identity is a precondition of one's being able to raise the question of identity, and cannot be something one establishes in attempting to answer that question.

That one remembers_w a past action is not (and could not be) one of the things one remembers_w about it, and neither is the fact that there is no branching in the M-type causal chain linking it with one's memory of it. And normally there is no set of remembered_w features of an action one remembers_w from the inside, or of the person who did the action, by which one identifies the action as one's own and the agent as oneself. If one has not identified a remembered person as oneself on the basis of his remembered_w features, then of course it cannot be the case that one has misidentified him on this basis. This is not to say that there is no basis on which one might misidentify a remembered_w person as oneself. If there can, logically, be remembering_w that is not remembering_s, then where one remembers_w an action from the inside one's judgment that one did the action will not be logically immune to error through misidentification in the sense defined in Sect. II—though given the contingent fact that all remembering_w is remembering_s, such judgments can be said to have a *de facto* immunity to error through misidentification. But the sort of error through misidentification to which a statement like "I saw a canary" is liable, if based on a memory_w from the inside, is utterly different from that to which a statement like "John saw a canary" is liable when based on a memory_w of the incident reported. If the making of the latter statement involves an error through misidentification, this will be because either (1) the speaker misidentified someone as John at the time the reported incident occurred, and retained this misidentification in memory, or (2) at some subsequent time, perhaps at the time of speaking, the speaker misidentified a remembered_w person as John on the basis of his remembered_w features. But if I remember_w from the inside someone seeing a canary, and am mistaken in thinking that person to have been myself, it is absurd to suppose that this mistake originated at the time at which the remembered_w seeing occurred. Nor, as I have said, will this be a misidentification based on the remembered_w features of the person who saw the canary. What could be the basis for a misidentification in this case is the mistaken belief that

there is no branching in the M-type causal chain linking one's memory with the past incident. But a misidentification on this basis, while logically possible, would be radically unlike the misidentifications that actually occur in the making of third person reports.

VII

Because I have taken seriously the possibility of worlds in which M-type causal chains sometimes branch, and thus the possibility of quasi-remembering (remembering_w) that is not remembering_s, I have had to qualify and weaken my initial claims about the "special access" people have to their own past histories. But if our concern is with the elucidation of our present concept of personal identity, and with personal identity as something that has a special sort of importance for us, then it is not clear that the possibility of such worlds, and the qualifications this requires, should be taken as seriously as I have taken them. For there is reason to think (1) that some of our concepts, perhaps including the concept of a person, would necessarily undergo significant modification in their application to such worlds, and (2) that in such worlds personal identity would not *matter* to people in quite the way it does in the actual world.

There are important connections between the concept of personal identity and the concepts of various "backward looking" and "forward looking" mental states. Thus the appropriate objects of remorse, and of a central sort of pride, are past actions done by the very person who is remorseful or proud, and the appropriate objects of fear and dread, and of delighted anticipation, are events which the subject of these emotions envisages as happening to himself. And intentions have as their "intentional objects" actions to be done by the very person who has the intention. It is difficult to see how the notion of a person could be applied, *with these conceptual connections remaining intact*, to a world in which M-type causal chains frequently branch, e.g., one in which persons frequently undergo fission. If I remember_w from the inside a cruel or deceitful action, am I to be relieved of all tendency to feel remorse if I discover that because

of fission someone else remembers_w it too? May I not feel proud of an action I remember_w from the inside even though I know that I am only one of several offshoots of the person who did it, and so cannot claim to be identical with him? Am I not to be afraid of horrible things I expect to happen to my future offshoots, and not to view with pleasant anticipation the delights that are in prospect for them? And is it to be impossible, or logically inappropriate, for me knowingly to form intentions, and make decisions and plans, which because of the prospect of immanent fission will have to be carried out by my offshoots rather than by me? To the extent that I can imagine such a world, I find it incredible to suppose that these questions must be answered in the affirmative. The prospect of immanent fission might not be appealing, but it seems highly implausible to suppose that the only rational attitude toward it would be that appropriate to the prospect of immanent death (for fission, unlike death, would be something "lived through"). It seems equally implausible to suppose that a person's concern for the well-being of his offshoots should be construed as altruism; surely this concern would, or at any rate could, be just like the self-interested concern each of us has for his own future well-being. Yet a negative answer to my rhetorical questions would suggest that either the concept of a person or such concepts as those of pride, remorse, fear, etc., would undergo significant modification in being applied to such a world.²⁷

A person's past history is the most important source of his knowledge of the world, but it is also an important source of his knowledge, and his conception, of himself; a person's "self-image," his conception of his own character, values, and potentialities, is determined in a considerable degree by the way in which he views his own past actions. And a person's future history is the primary focus of his desires, hopes, and fears.²⁸ If these remarks do not express truths about the concept of personal identity, they at least express truths about the *importance* of this concept in our conceptual scheme, or in our "form of life." It seems plausible to suppose that in a world in which fission was common personal identity would not have this sort of

²⁷ On this and related questions, see my exchange with Chisholm in *Perception and Personal Identity*, *op. cit.*

²⁸ This is not to deny the possibility or occurrence of unselfish attitudes and emotions. Even the most unselfish man, who is willing to suffer that others may prosper, does not and cannot regard the pleasures and pains that are in prospect for him in the same light as he regards those that are in prospect for others. He may submit to torture, but he would hardly be human if he could regularly view his own future sufferings with the same detachment (which is not indifference) as he views the future sufferings of others.

importance. Roughly speaking, the portion of past history that would matter to a person in this special way would be that which it is possible for him to remember_w, and not merely that which it is possible for him to remember_s. And the focus of people's "self-interested" attitudes and emotions would be the future histories of their offshoots, and of their offshoots' offshoots, and so on, as well as their own future histories. In the actual world it is true both that (1) remembering_w is always remembering_s (and thus that there is special access in the strong sense characterized in Sect. I), and that (2) the primary focus of a person's "self-interested" attitudes and emotions is his own past and future history. It is surely no accident that (1) and (2) go together.²⁹

The Rockefeller University

Received December 9, 1969

²⁹ This is a considerably revised version of a paper which was read at a conference on "The Concept of a Person" at the University of Michigan in November 1967, and at the University of British Columbia and the University of Saskatchewan at Saskatoon in the Spring of 1969. I am grateful to Harry Frankfurt, Robert Nozick, and Michael Slote for criticisms of the earlier versions of the paper.