

MEMORY AND PERSONAL IDENTITY

P. M. McGoldrick
Lake Forest College

If I remember seeing the first Sputnik 23 years ago, hearing the 1:00 P.M. radio news five minutes ago, and many other things in between, does that imply that I am now the same person I was then?

What does the question ask? It asks whether an entity which is the percipient of my present experiences is numerically identical to the entity which was the percipient in a series of remembered experiences. We should note the distinction between 'percipient *of* my memory experiences' and 'percipient *in* my memory experiences.' The former refers to the percipient of the memory itself and is clearly the present subject of consciousness, while the latter refers to the percipient of that experience which is chronicled in the memory record. Only in the latter case does it make sense to ask whether a relationship of numerical identity holds between the two entities, since it is only here that it could seriously be doubted.

How do we normally decide whether some entity which exists at one time, is the same entity as that which exists at some other time? Taking a clue from Mackie,¹ we might suggest that all identity statements are necessarily relative to some sortal notion, for in no other way can we specify the criteria relevant for judging of a thing that it is the same as that which existed at some earlier time. In addition, the entity must exhibit a continuous history under a single sortal classification, throughout the period for which a numerical identity relationship is claimed. Thus, the tree growing outside my window is the same tree as the one that was there two years ago, but the molecule cluster occupying that very space is not the same molecule cluster which was there two years ago. This is because through there has been a continuous 'tree' history on that spot for the past two years, the particular molecular components have periodically perished to be replaced by others.

If we attempt to treat the problem of personal identity as a subspecies of the general problem of entity identity across time, we encounter the following difficulty: the criterion on the basis of which we decide whether or not an entity has had a continuous history under the sortal notion 'Person X', differs from first person to third person judgments. Deciding whether the Lee Harvey Oswald who left the United States to visit the U.S.S.R. is the same Lee Harvey Oswald who assassinated Kennedy is achieved by paying close attention to his physical appearance, by checking his fingerprints and dental records, and by comparing his account of his personal history with the known facts. But the man who assassinated Kennedy knows whether or not he is the same Lee Harvey Oswald who visited the U.S.S.R. in a quite

different way. Independently of any appeal to publically observable evidence, he simply remembers going, or not going, to the U.S.S.R., and so on, and hence knows that he is, or is not, the same person.

The existence of two distinct, and apparently independent, criteria of personal identity presents an obvious problem; it is quite conceivable, though admittedly instances of it are rare, that a judgment made on the basis of one criterion, e.g., the publicly observable body criterion, could come into conflict with a judgment made on the basis of the other, e.g., the memory criterion. One way of dealing with this might be to argue that one or the other criterion has, as a matter of fact, priority over the other. For example, it might be argued that an individual is in a more privileged position to know whether or not he is the same person as the one who existed at some earlier time, and that, therefore, priority ought to be given to first person judgments and the memory criterion upon which they are based. But, as a matter of fact, we do not give priority to one or the other criterion. Both are usually appealed to, and usually they corroborate one another. In those cases where they conflict, we normally withhold judgment because we don't know what to say.

In this paper, I want to argue that priority ought to be given to one criterion, the publically observable bodily criterion, not because this reflects practice, but because the memory criterion provides inadequate evidence for an identity claim. In other words, I want to argue that the fact I remember seeing the first Sputnik, hearing the 1:00 P.M. radio news, and many other things in between, does not imply that I am now the same person I was then.

Some of the most convincing considerations offered in favour of the adequacy of the memory criterion, have been put forward by philosophers who draw on Kant's argument for the transcendental Unity of Apperception in the *Critique of Pure Reason*.² One such philosopher, H. J. Paton, gives a rather succinct version of the argument in his article, *Self-Identity*.³ Paton considers a man who, through a limited temporal duration, hears a clock strike twelve. A necessary condition of his hearing all twelve strokes is that he hear each one of them. This seems fairly obvious. It is equally necessary that the strokes which occur at the earlier time be retained in the man's consciousness at the time when the last strokes are heard. Kant characterizes this condition as "retention in imagination," but Paton claims that this is an unnecessary complication when "all we require is memory." Again, this seems fairly obvious. A man suffering from a persistent amnesia which erased the memory of his experiences as fast as he had them would be incapable of hearing a clock strike twelve; by the time the last stroke sounded he would have forgotten that he had heard the first. A third and final condition is that he recognize that that which is retained in memory is the same as that which he heard at the earlier moment. In other words, he

must recognize that what he remembers is 'the first stroke of the clock', 'the second stroke of the clock', and so on. For, if he fails to recognize that the items in his memory come under the sortal 'first clock stroke', 'second clock stroke', etc., then he will, as Paton says, fail "to grasp the whole series as a unit, or unity, or whole, e.g., as 12 strokes" (1929, p. 315).

Paton holds that the above considerations are sufficient to establish the thesis that the apprehension of a series of experiences in time implies a single conscious subject of those experiences. For, "if the twelve strokes are heard by 12 different selves, there is no hearing, or at any rate no knowing, of the twelve strokes" (1929, p. 315). That is to say, if a person remembers a series of experiences, then necessarily he is one and the same person who experienced each and every member of the series; if he were not, they could not have formed a series, but would be completely disparate occurrences.

Surprisingly enough, this inference, which so many philosophers are prepared to make on Kants' behalf, does not appear to be one he himself wanted to make. In the section on the Paralogisms in the *Critique of Pure Reason*, he says:

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, and yet we may ascribe to it the same-sounding "I" which in every different state, even in one involving change of the [thinking] subject, might still retain the thought of the preceding subject and so hand it over to the subsequent subject.⁴

Mere awareness of a unified sequence of experiences in memory is insufficient to ensure that the self-same subject of consciousness is the percipient of each and every member of the series. Each may have been perceived by a distinct subject, which transmitted the record of its experiences to a subsequent subject, thereby generating a memory record of sequential experiences. In a footnote, Kant offers as an illustration the example of an elastic ball in motion which comes in contact with a second elastic ball. By impinging upon it, the first ball transmits its motion to the second. In the same way, he suggests, we might postulate subjects of consciousness "such that the one communicates to the other representations together with the consciousness of them. . . . The last substance would then be conscious of all the states of the previously changed substances, as being its own states, because they would have been transferred to it together with consciousness of them. And yet it would not have been one and the same person in all these states."⁵

Let us put the point in a way which is more within the realm of scientific

possibility. Recent developments in neurobiology show that all the experiences of a percipient which are subsequently remembered are first encoded into unique pathways or circuits of neurons in the nerve fibers of specific parts of the cortex. By artificially stimulating these circuits with electrodes, the encoded memory is brought back into consciousness. Chemically treating the brain tissue causes these neural circuits to floresce when illuminated with ultraviolet radiation, and as a result their pathways can be mapped out in the form of micrographs.⁶ Now imagine that we had a way of artificially imprinting neural circuits onto a cortex, just as we imprint integrated circuits onto silicon microchips. It would then be possible to transmit particular memory sequences from one brain to another. Parents, wishing perhaps that their offspring be experienced in the ways of the world from an early age, might have their particular memory sequence transmitted to their children, mothers to daughters and fathers to sons. If this process were to continue over a number of generations, it would become commonplace for people to have memories which stretched back for hundreds of years. Thus, a man might remember the battle of Waterloo, the General Strike, and World War Two; as a consequence, it would appear to him that he was one and the same person who experienced each of these events, though as a matter of fact they would have been experienced by a number of different subjects of consciousness. This surely is what Kant had in mind when he distinguished between the empirical and transcendental self, and insisted that what I think of as myself is a mere appearance, based upon remembered states, and therefore, an empirical self. Since all knowledge is based upon experience, the transcendental self, which is the real subject of thought, is unknown, and therefore, it is unknown whether it be a single self or a diverse cluster of selves.

It might be objected that the man in our example cannot truly be said to remember, e.g., the battle of Waterloo, since it is false that he was alive in 1815. A memory claim, it might be said, is more like a knowledge claim than a belief claim; when we discover that it is false, either because the "remembered" event never occurred or because the agent who claims to remember it was elsewhere at the time, we rescind the claim. If I say I remember seeing the first Sputnik, but evidence shows I was in the Southern Hemisphere at the time, I am obliged to say that I was mistaken.

This is the sort of point which John Williamson wants to make in his article, "The Persistence of Memory."⁷ That personal identity endures, he argues, is logically entailed by *the truth* of those propositions which express memory claims. Remembering that something is the case implies knowing, at the time of remembering, that it is so. Why? Because one cannot remember what one has never known; hence, whatever is remembered must have been known at some point by the rememberer. Moreover, since it is also true that one cannot remember what one has forgotten, whatever is

remembered must be known at the actual time of remembering. From this it follows that remembering seeing the first Sputnik 23 years ago and hearing the 1:00 P.M. radio news five minutes ago entails knowing that I saw the first Sputnik and heard the 1:00 P.M. news.

Now, since one cannot know what is false, any proposition which one can be said to know, including first person propositions, must be true. Thus, if I know that I saw the first Sputnik, then it must be true. Thus, if I know that I saw the first Sputnik, then it must be true that I saw it. But if it is *true* that I saw it, then I must be the same person as the person who saw it, since the predicate "having seen the first Sputnik" is *true* of the present knowing subject. Thus, knowing past-tense first person propositions necessarily entails an identity across time between the present subject of consciousness and the subject of those propositions. And since remembering implies knowing at the time of remembering, it follows that rememberings expressed in first person propositions likewise entail an identity across time between the two subjects of consciousness.

To evaluate this argument we need to ask what we mean when we say of someone that it is true that he remembers an event. In other words, we need to ask under what conditions one can be said truly to remember something. Clearly, the mere fact that a man claims to remember an event, however sincere he may be, is insufficient. We are all too familiar with the phenomenon of mis-remembering. Memory claims must be corroborated before we can be certain that they are true. And corroboration is *usually* provided by evidence which establishes that there was a continuous body history from the time of the remembered event to the present. Thus, we are prepared to say of someone who claims to remember the Aurora Borealis as a child, that he truly remembers seeing it if parents and relatives, etc., testify to this fact, but not otherwise.

In some rare cases, we might be prepared to consider waving the need for a continuous body history. Consider the following case which is cited by C. J. Ducasse in his book, *Nature, Mind and Death*.⁸ Katsugoro, a young Japanese boy, claimed to remember a previous existence a few years before his birth. He claimed his name was Tozo, that he was the son of a farmer called Kynbei who lived with his wife in the village of Hodokubo. In addition, he recounted how his father had died, how his mother remarried a man called Hanshiro, and how he himself died a year later from smallpox. When, eventually, he was taken to a village called Hodokubo, and despite the fact that neither he nor his family had ever been there before, he immediately led the way to a house which, it transpired, was occupied by the very people whom he had claimed were his parents. Moreover, not only did he recognize many buildings in the village, he also correctly pointed out a shop and a tree which had not been there when Tozo lived in the village.

This case is interesting because, although the events which occurred

subsequent to his visit to the village of Hodokubo are insufficient to establish the truth of Katsugoro's memory claims, and are, therefore, insufficient to establish that he is one and the same person as Tozo, the very fact that his story is corroborated by external evidence makes us hesitate before dismissing it as mere fantasy or imagination. This suggests that corroboration by *external* evidence of some kind or other is the key feature which determines whether a memory claim is accepted as true. Thus, *by itself* the memory claim provides no justification for its own veracity. But if this is so, then a memory claim alone is insufficient to establish that the percipient in the remembered experience, and the rememberer, are one and the same person. In other words, memory is not a sufficient condition of personal identity.

Nor is it a necessary condition. Although an amnesiac lacks a memory record, he is not thereby robbed of his personal identity. The mere fact that he cannot remember the past events which constitute his particular life history does *not* lead us to conclude that, therefore, he is not the same person as the one who experienced those now forgotten events. We may say, "poor fellow, he cannot remember who he is," but we are quite confident that he is the same person as the one who experienced such and such a particular life history, *if* there is sufficient external evidence to confirm this. Memory, therefore, is neither a sufficient nor a necessary condition of personal identity. It is not sufficient because even when it is present we may still want to withhold the identity claim, as in the Katsugoro case. And it is not necessary because even in its absence we are unwilling to deny identity, as the case of the amnesiac clearly illustrates.

But if memory is neither a necessary nor a sufficient condition of personal identity, on what basis is the issue settled? This question can be answered in a fairly simple way if we examine the reasons why we are reluctant to allow identity in the Katsugoro case and why we are reluctant to disallow it in the amnesia case. In the former, identity is withheld because a continuous body history is lacking, while in the latter case identity is retained precisely because a continuous body history is present. Clearly, therefore, a continuous history under the sortal notion 'body' is not only necessary, but also sufficient to establish personal identity. This may seem a rather strange conclusion to come to, but it is less so when we consider that establishing bodily continuity is the only means available whereby we can verify memory claims. And this is so regardless of whether those memory claims are our own or another's. Thus, the fact that I remember seeing the first Sputnik 23 years ago and hearing the 1:00 P.M. radio news five minutes ago, as well as many other things in between, does *not* entail that I am the same person now as I was then. The 'I' may, as Kant suggests, refer to a variety of different selves, each of which transmits its memory record to a subsequent self. Moreover, when we make the judgement that there is

identity of personhood across time, we *mean* thereby not that a single self has persisted throughout the temporal duration but that a single body has.

NOTES

1. J. L. Mackie, *Problems from Locke*, Oxford (1976).
2. I. Kant, *Critique of Pure Reason*, trans. Kemp-Smith, pp. 135-61.
3. H. J. Paton, "Self-Identity," *Mind*, 38 (1929).
4. I. Kant, *op. cit.*, p. 342.
5. *Ibid.*, p. 342.
6. E. R. Kendal, "Small Systems of Neurons," *Scientific American*, Sept., 1979.
7. J. Williamson, "The Persistence of Memory," *Proceedings of the Aristotelean Society*, Nov., 1979.
8. C. J. Ducasse, *Nature, Mind and Death* (Open Court: 1951), pp. 492-93.