Macalester Journal of Philosophy

Volume 6
Issue 1 *Spring, 1996*

Article 5

2-17-2011

The Color of Memory

Sebastian Lange
Macalester College

Sebastian Lange

"The Color Of Memory"

> "...[C]onsciousness of what is past thus ascertains our personal identity...." (Butler, 100)

Introduction

Many philosophers have not been satisfied with the attempts to construct an account of personal identity on physicalist or Cartesian grounds. Instead, many have chosen memory as the centerpiece of their theory. Strangely, however, works about personal identity rarely offer satisfactory accounts of memory. Opponents and defendants of personal identity seem to neglect the crucial center of what might constitute or inhibit personal identity: the character of memory.

This paper has two main aims: to introduce a theory of memory based on empirical evidence, and to illustrate the effect such a theory should have on personal identity theories based on memory. I will develop my own account of memory as consisting of four key points: 1) Memories are encoded in the brain; 2) Memory is constantly "mutating" on two plains--*vertical* (there are immediate factors that influence the memory recall), and *horizontal* (there is a long-term alteration of memory over time); 3) There is no "unique, original memory"; 4) It is not possible to recall a memory in exactly the same way twice.

Following from these I will suggest that an identity relationship based on memory will be difficult to construct. Such a theory will have to take into account the severe changes that take place with the material that the connections into the past are made out of: mutating information and non-self-contained memories. A personal identity theory based on remembrance will not be able to ignore the continually changing "color of memory."

Locke's Account of Personal Identity

One of the first philosophers to give a positive account of personal identity based on memory was John Locke.[1] In *Of Identity and Diversity,* Locke argues that it is exclusively memory which constitutes our personal identity. We are born as a *tabula rasa* later to be "filled" by experiences, which determine who and what we are.

In order for an identity relationship to hold over time, Locke deemed it necessary to make memory a static entity:

---

[1] Hobbes and Hume also treated memory as the most likely centerpiece constituting personal identity. They were, however, much more reluctant to conclude that memory actually does constitute personal identity.

> ...For as far as any intelligent being can repeat the idea of any past action with the same consciousness it had of it at first, and with the same consciousness it has of any present action; so far it is the same personal self. For it is by the consciousness it has of its present thoughts and actions, that it is self to itself now, and so will be the same self, as far as the same consciousness can extend to actions past or to come... (Locke, quoted in Perry, 40).

Locke seems to argue that remembering an event X means regaining the same content of consciousness of event X. If this were true, we would indeed have an uncomplicated identity relationship. A person is identical with itself at any instance because the content of her consciousness is identical with itself at a given moment. Thus, if the content of consciousness, which is identical with itself at any instance, could be "filled" with the content of consciousness that the agent had when experiencing X in the past, we can say that the person now and the person experiencing X are identical with each other; they both share the same content of consciousness.[2] This means a transitive identity relationship could be constructed, viz., Person A=C and Person B=C; therefore, Person A=Person B, where C stands for a same consciousness. But do not our most ordinary everyday experiences seem to disprove this claim?

Objection

In the following sections I will venture to support my claims about memory by psychological and scientific studies, and hence try to place the notion of memory on a factual basis. However, before proceeding, I must disclaim an objection that could undermine the legitimacy of the whole paper.

It might be objected that science currently has no conclusive knowledge about memory. Some of the connections between brain activity and memory remain a mystery. Therefore, uncertainty of the processes of memory, it is objected, will render doubtful any empirical evidence based on a highly incomplete, inconclusive body of scientific knowledge. This can mean either that the evidence I bring forward is in part only a temporary scientific standpoint, or that it is completely mistaken to take such empirical data as a part of a philosophical argument at all.

The first point I acknowledge. There will be, without any doubt, countless improvements of the current scientific viewpoint on memory; better experiences will be designed, new instruments devised, etc. Yet, I maintain that even with the current data, a general trend is to be seen--the changeability and mutability of memory.

Secondly, it is untrue that uncertainty about a mechanism which underlies accessible effects should render all claims about this mechanism useless. Let me give an analogy. If a physicist came up to me today and said that all stones on this planet should be perceived as suspended in midair, I would have two ways of reacting: either I condole myself, and all the people whom I meet who perceive just as I do that stones

---

[2] Locke's meaning of the word "consciousness" differs somewhat from the modern, ordinary conception of consciousness.

seem to fall, because we apparently would be living in a life-long delusion; or I doubt the story of the physicist.

Here the second reaction seems the more credible, because if the physicist's theory failed to explain just why humans perceive stones that fall as falling, so much the worse for his theory. The physicist maintained that one ought to perceive a certain characteristic, not that behind ordinary perception, whatever it may be, stones display a certain behaviour. Thus, one has good grounds for making a fair judgement about his theory because one can base such judgement directly on one's own perception. One did not need to know anything at all about the underlying mechanisms of gravity to make a useful claim about this mechanism: it does not give rise to the effect that all stones you perceive should be suspended in midair. One also did not need to know anything about the physicist's theory besides his one claim to make a useful remark about it. In the same line of reasoning, I maintain that we can make useful statements about the nature of memory without knowing all parts of the intricate workings of memory (if that should ever be possible).

Such a useful claim is at best an inductive truth; that, I acknowledge. Let us now turn to some details about memory.

Different types of memory

The classification of memory has for long been a dispute between experts in psychology and philosophy. It would be an exaggeration to proclaim that a detailed agreement has been found; it has not. The more artificial and arbitrary the definition of principles and observations relating to a term, the more likely it is that such a term will be disputed or even rejected out of hand. However, the classification (stemming from Artificial Intelligence research in the 1970s) of memory into *two* main types (procedural and declarative) has drawn relatively little critique.

Procedural memory is the memory of *how* something was performed. Sometimes this type of memory is also referred to as skill, or "habit memory." For example, 15 years ago I learned the skill to ride a bike. Afterwards I did not have to concentrate on when I had to push what pedal at which moment. Some time later I even learned to ride without having my hands on the handlebars. Only a triggering command was necessary to perform such action, not a concentrated effort to direct every minute movement.

Another example for a procedural memory is the concert pianist. When playing Rachmaninoff's variation on Paganini's Twenty-Four Caprices, the pianist does not concentrate and visualize every one of the complex movements of his hands. Rather, many pianists claim they were scarcely aware of all the complex hand movements that went on when performing the piece. Years and years of training enabled such a person to direct his playing. In fact, it would be impossible to play most piano pieces if every hand movement had to be consciously directed; too many tasks need to be performed in parallel.

Yet, to know what a piano is, to know what my experience of learning to ride the bike was like (e.g., that I bruised my knees)--in fact, to know what meaning anything has--belongs to the second type of memory. *Declarative memory* are things

that can be consciously recalled. The declarative memory itself can be divided into a semantic and episodic memory.

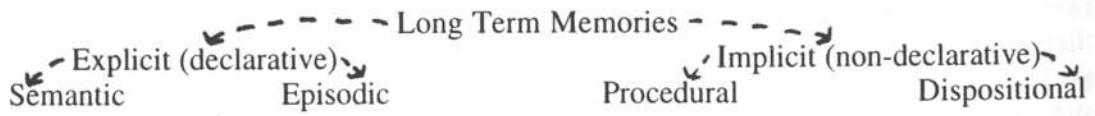
Semantic memory is knowledge independent of one's own life history (for example, the knowledge that, despite the disclaimers of some, the Holocaust took place). Episodic memory, on the other hand, is the memory of events in one's own life history.[3] The latter (and therefore, the whole group of declarative memories, in contrast with the group of procedural memories) seems of great importance for personal identity theories based on memories, because they can enable the construction of a personal narrative which can be consciously recalled. This, in turn, can enable an agent to claim to be identical with a certain person in the past because she takes memories that are available to conscious retrieval as evidence for such an identity. For these reasons I will make much use of examples dealing with declarative memories.[4]

At first glance, it may be surprising that memory is not a unified system. Different attributes of a present experience will influence a different type of memory (procedural or declarative). For example, when playing a service in tennis, I do have a feeling of what it is like to play that service (throwing the little yellow ball into the air and fiercely swinging my racket from behind my back to hit the ball). At the same time, I notice the old oaks behind the fence of the tennis court and then focus on the location to whence I want to play my service. These two aspects do not seem to be divided when I experience them, yet in memory they can be divided into procedural and declarative memories. My experience of what a service feels like might contribute to an increase of my skill in making good services. Yet, later I might only consciously recall a picture of the old oaks behind the fence as a part of my declarative memories. I might not consciously recall the feeling of *what* the swinging of the tennis rack felt like. However, the feeling of what the swinging of the racket felt like at the time of playing might have contributed to my procedural memory. This memory, however, could not be consciously recalled and therefore

---

[3] Such phenomenological systematization of memory (Rose, 120) might still lend itself to a dualistic account including the observed phenomena of memory, since they are not tied in this theory to the working of the brain.

[4] One could make a brief chart of the different memory types as follows:



(after Squire and Schacter, in Myers, 270)

Here non-declarative memory is not only consistent of procedural memory, but also of dispositional memory. Dispositional memory is, according to Squire and Schacter, classical and operant conditioning.

easily verified, whereas a conscious (declarative) memory could, even though both memories resulted from the same, experiential unified experience.

This distinction between the two main forms of memory (declarative and procedural) can even go so far that one could, as a result of brain damage, entirely forget "...that a two-wheeled object which one can sit on pedal and move about..." is actually a bicycle, and yet still be able and know how to take off and ride for five hours on this machine (Rose, 120). Such a distinction seems to be supported by a Canadian patient simply known to the scientific world under the initials H.M.

Because of strong epileptic tendencies, significant areas of H.M.'s brain were removed in an operation in 1953 (Rose, 124). After his surgery, H.M. was unable to remember persons, names and faces. Every moment seemed as if awakening from a dream. In other words, as a result of the surgery, his episodic memory of events was from then on inaccessible. Yet he could be trained a skill, and even when proclaiming every time when he was trained that it was the first time he performed the presented task, H.M. consequently improved on what he did. (For example, he was trained to play the game of building the so-called towers of Hanoi.) And here we arrive at the first point to notice: the *mutability* of memory.

> ...The inescapable conclusion is that procedural and declarative memory are not merely localized, but localized to different regions of the brain, so that the one, declarative, can be lost, whilst the other, procedural, is spared...it seems extremely difficult to lose procedural memory and relatively easy to lose declarative memory... (Rose, 128).

This means that the information derived from a specific event is stored relatively independently: so independently, in fact, that one type of information (declarative memory) seems more apt to alteration and loss than does the other (procedural memory).

In conclusion, a memory of an event does not seem to be represented in one solid block that can be recalled fully. A categorization of memory into procedural and declarative thereby speaks against Locke's notion of a consciousness of an event past as if it were present. The information derived from a past event is stored in two different memory systems. Yet, when an event is presently experienced, the feeling about the event, as well as the explicit thoughts and mental images, are unified. As the example of H.M. shows, both memory types are not necessarily linked in a recall. The distinction between declarative and procedural memory is a good indicator against the notion of a possible unified recall of a situation exactly as it appeared to the subject in the past. A Lockean might now limit the "same consciousness" simply to declarative memories, making them the cornerstone of an account of personal

identity. Yet, as will become clear in the next sections, even the declarative memory lacks the static stability sought for.[5]

The mood dependency of memory

Let us now turn to a more specific example of declarative memory recall. In recent studies (Seidlitz and Diener), an interesting relationship between the mood of an individual and the type of memory recall was uncovered.

First of all, it has to be clear what I mean by the term "mood." In the everyday use the word "mood" mostly refers to a momentary disposition of an agent.[6] For example, at any given moment I may be sad, happy, delighted, angry or aggressive. I will continue to use the word "mood" in the sense of *momentary emotional disposition (MED).* In their study about the relationship between the disposition of an agent and the agent's memory recalls, Seidlitz and Diener distinguish yet another form of disposition: *the subjective well being (SWB).*

> ...SWB consists of two components, long-term happiness and satisfaction with life.... Happiness refers to average levels of positive and negative affect considered over a long time period... (Seidlitz and Diener, 654).

In this sense, SWB seems to be a long-term outlook on life, whereas the word "mood" in the sense of MED has a more temporary character. For example, I can feel angry at this moment (e.g., in a mood, MED) but still have a happy, positive SWB.

In a series of tests Seidlitz and Diener asked a number of subjects to memorize a list that contained words with neutral connotations (as far as this can be reached),[7] negative connotations[8] and positive connotations. Afterwards, an essay that the subject had formerly written was critiqued--either in a neutral way (no special remarks), in a rewarding way ("your essay was the best among all tested persons") or in an negative way (e.g., dramatically tearing up the writing, proclaiming the subject

---

5 The aforementioned discovery of the dichotomy between declarative and procedural memory also casts doubts on Parfit's try to amend a Lockean account of personal identity; more on this in my closing paragraphs.

6 Agent is commonly replaced by "person"; as my last paragraph will try to suggest, we should be careful using this term in philosophical writing. Since I do not want to bring about confusion using a word in a common sense and later redefining it in a philosophical sense, I shall refer to a human being as "agent."

7 Since, of course, the connotation of a word can differ from agent to agent, more than a very few agents had to be tested in order to make sure that some general trend can be concluded.

8 For example, words with very pejorative meanings that I would never allow to appear in print **.

either does not try or is utterly incapable of formulating a valuable idea). In this way, a momentary mood of either happiness or anger was induced in the subject. The angry subject would remember clearly more negative words than the happy subject. It therefore appears that the character of one's memory is dependent on one's immediate mood.

If one widens the border of immediate mood to an immediate disposition (including physical well-being), this claim can even be further supported: it was shown that immediate "...bodily discomfort did facilitate negative memories...."[9] Another study at Yale suggests that under certain circumstances,[10] when a person is put in a stressful condition, "powerful flashbacks" occur.[11] Extraordinary memory events such as a "powerful flashback" seem to be in some way causally dependent upon the immediate disposition of an agent.

I therefore conclude that the type of memory, the selection of what is remembered (event-feature selection) and in what way it is remembered (if a memory appears as a positive, neutral or negative event) is influenced by a subject's MED and SWB as well as by her/his physical well-being.

In the same study Seidlitz and Diener point to a connection between an agent's subjective well-being and memory. It is not clear at what stage memories becomes "biased" (interpreted either in a good or bad way when already encoded in memory [Seidlitz, 659]); however, an overall happy agent will tend to interpret an event that is not clearly either good or bad as a positive event and recall it as such.

In this way the subjective well-being is like a dynamic system which comes to reinforce its own character constantly by interpreting memories in the relevant way (as good/bad/neutral), after which these memories themselves become reinforcements of the respective SWB, which in turn interprets the next memories even more distinctively. Such constant self-construction might enable us to be able to assign some character traits to an agent over a longer period of time (months and years). However, SWB is itself bound to change. An interpretation of an event as reinforcement of the current SWB may itself change the outlook on life slightly, thus changing the SWB itself. This means that it also changes the interpretation of memories itself (since, as Seidlitz and Diener showed, memories seem to be interpreted and influenced by an agent's SWB).

Also, the overall happiness with life (SWB) might be disrupted by an event, that, for example, causes such a great distress that it can not be interpreted as good. This pushes the SWB in a different direction. Many small, negative events might also amass, and can gradually change the outlook of an agent on her/his life, and thus "her/his satisfaction with life." Such an alteration would also mean a change of the SWB, according to which other memories and events are subsequently interpreted in a slightly different way.

---

9 Taylor, 876, about a study by Berkowitz in 1990.

10 When noradrenalin--a neurotransmitter associated with alertness and stress--is released in the brain (Kandel and Kandel, 36).

11 Here, explicitly tested on Vietnam veterans (Kandel and Kandel, 34).

In summary, we have already encountered some important building blocks for a theory of memory. Memory is mutable in two directions: over the long term (old memories are re-interpreted by a change in SWB), and in an immediate fashion (memory recalls are influenced by immediate factors through immediate mood (MED) or physical (dis)comfort).

The usage of SWB as a foundation for an account of personal identity seems to lead to great difficulties. "Happiness and satisfaction with life," which constitute subjective well-being, are very vague notions. Even though a subject can credibly report about her/his being satisfied or happy in one way or the other, it is nevertheless extremely difficult to account for just "how," "in what way" or "how much" one is happy or satisfied (the problem that some forms of utilitarianism run into). It seems exceedingly difficult to quantify a long-term outlook on life. This means that a distinction between the SWB of an agent at time t1 and the SWB of an agent at time t2 cannot be differentiated in quantitative terms. For example, the agent at time t1 and time t2 might say that "I feel content with my life," but there is little way of telling whether the form and strength of being content at t1 is really similar to being content at time t2. This makes it difficult to construct an identity relationship based on SWB for the agent at t1 and t2, because we may not be able to make sure that the SWB at t1 and t2 are the same.

But even if we assume that we could give enough characteristics of the subjective well-being at two points in time to render them identical, the nature of SWB seems to preclude such a possibility: As has been shown in Seidlitz and Diener's experiments on the relationship of memory and SWB, the SWB is bound to change in the form of reinforcement or inhibition of its current character. Yet, it has to be remembered that what we are looking for in a theory of personal identity is *something* (let me call it C) that endures unchanged, unaltered in time, so that for a person A at time t1, C is a member of A, and for a person B at time t2, C is also a member of B. The SWB seems bound to change in time, and thus, no identity relationship of the above kind might hold. Therefore, the subjective well-being, a long-term outlook on life, can hardly be the "Holy Grail" which defenders of the existence of personal identity would want to pick out as the feature that produces personal identity. It is, again, important to bear in mind that the more facts that are uncovered about memory, the more unstable and mutable memory seems to be: It has been observed that memory is influenced by immediate as well as long-term factors. This leads to the consideration that memory cannot be a static entity over the short, as well as the long, run.

Repressed memories

The idea of repressed memories became popular through the works of Freud, relying on Breuer's discovery about the influence that a traumatic experience can have on an agent (Stevenson, 76). Freud's system of psychoanalysis is based on the presumptions that repressed memories exist, and that such memories exert influence on an agent. Such influence can be the source of a patient's mental suffering (e.g., neurosis). A gist of Freud's theory about repressed memories runs as follows:

> ...In a situation of extreme mental conflict, where a person experiences an instinctual impulse which is sharply incompatible with the standards he feels he must adhere to, it is possible for him to put it out of consciousness, to flee from it, to pretend that it does not exist. So repression is one of the "defense mechanisms" by which a person attempts to avoid inner conflict... (Stevenson, 76).

Even though Freud's tripartite structure of the human mind has been criticized severely, his work made acceptable the idea of existing unconscious mental states.[12] The fact that people such as Daniel Dennett arrive at the same notion from completely different starting points lends this theory some further support.

In current times the news of repressed memories flare through the media because of the many cases of child abuse trials that rely on previously repressed, and now recovered, memories. The validity of such recovered memories was questioned and has created the need for more research in this area. Judith Hermann at Harvard found that the repression of memories in abuse cases is not uncommon. Of 53 women attending incest survivor groups, almost two-thirds reported to have had memory lapses at some point after the abuse occurred: "...Most clinical psychologists believe that children can learn to block memories as a survival mechanism..." (Kandel and Kandel, 37).

Up to now there has been no definite answer as to how repression might be accounted for biologically. What is interesting is that when repressed memories return, most victims have a feeling of familiarity, before they are able to recall the forgotten episode in its details. The memory is not a descriptive one at that point, but only a strange feeling of familiarity can be experienced. The individual at such stage cannot recall any episodes about the repressed event. A descriptive memory (i.e., a detail of the repressed episode) only follows at a later stage, if at all.

What makes repressed memories important in my argument is that they demonstrate very graphically how over a period of time the character of a memory can alter. First memory was a non-retrievable, but influential, memory. Then the repressed memory "surfaced," and the agent became aware of his/her past experiences (assuming that they were not a product of a psychologist's suggestive power). The memories are now amongst the set of memories that can be recalled in some way (preconscious, as Freud would say). The character of the memory has clearly changed over time.

Of course, child abuse is only a very special case, and investigations need to be done to see if repression of memory occurs in other areas. The door is even open to wonder whether unfavorable experiences of kinds other than child abuse can be subject to equal repression. Many cases of neurosis seem to have supported the theory that repression occurs not only in abuse situations. But if this is true, we

---

[12] In fact, in Kandel and Kandel's article, one finds a definition for repressed memories similar to that of Freud's.

might be subject to frequent recovery of memories without realizing it. However, I will not speculate further to find some factual basis for a theory of memory.

Memory can change its character over time; it is mutable. The change can produce memories which are influential, yet cannot be consciously recalled. At another time, with the right incentives, the agent can become reacquainted with his memory.

Forgetfulness

If one thinks that forgetfulness is a shame, it is worth looking at the case of the Russian, Sherevski; he had difficulties forgetting (Rose, 100). Sherevski was able to remember an elaborate (but senseless) mathematical formula after 15 years of not being reminded of it. At first glance, such a characteristic seems wonderful, but "...it was hard for [Sherevski] to make normal human relationships because he found it difficult to merge his recollections of people..." (Rose, 102). His memories are like snapshots, fixed and frozen in time.

The case of Sherevski sheds some light on the way we normally remember. It seems quite difficult for us even to imagine to remember the way Sherevski does (for a telling characterization of such a person, see J.L. Borges' story, "Funes the Memorious"). Forgetting seems a normal attribute of adult life. The fact that we recognize the case of Sherevski's memory as extraordinary should make us suspicious about a claim that demands the remembrance of every detail of a past experience, especially if our personal identity is explained by such a way of remembrance.

It is interesting to note that psychology affords a recognized distinction between the way adults normally seem to remember and the way Sherevski seemed to remember. This indicates that a change of type of memory over time seems to be a recognized phenomena. *Eidetic memory*[13] is the name of the detailed, snapshot-like memory exemplified by Sherevski. It occurs very rarely in adults; yet, when elementary school children were examined, it was discovered that about half of them do have eidetic memory. In the 1960s and '70s U.S. scientists confirmed this trend (a little below 50% of the children) for elementary schoolchildren, independent of sex, race or social origin (Rose, 103). Such observation seems to coincide with folk proverbs about the "special kind of fantasy of young children" or could be part of the explanation why it often seems very easy to learn a new language very fast for young children.

It was found that after puberty, the cases of occurrence of eidetic memory decreases rapidly and is minimal in adulthood: "...The transition from childhood to adult memory is dramatic, from imaged and timeless to linear and time-bound..." (Rose, 106). It follows from these observations that the character of "...memory itself is a developmental process..." (Rose, 104). (For example, in individuals "...between nine and ten [years of age], the quality of memory has changed, whereas [in individuals] between forty-nine and fifty, it is stable..." [Rose, 104].)

---

13 Eidetic memories are, of course, a subgroup of what I have called declarative memories.

If one accepts the preponderance of eidetic memory in young children and recognizes at the same time the peculiarity of Sherevsky's case, then it seems to follow that the fundamental way we do remember changes over time in accordance to the stages of growing up: At a young age, we possess the one mode of memory (eidetic); at another stage, the other (adult memory). Such changes might be gradual and unnoticed, but if the two extreme-points (late adulthood and early childhood) are viewed together, the character of memory is very different in each case. As cited, the studies (on Sherevski and schoolchildren) about the type of memory suggest that alteration of the way we remember is not an improbable contingency. On the contrary, it seems to be the case that a change of mode of remembrance is an ordinary part of growing up. But this would mean that it is very improbable that memory (at least in early stages of life) can be a static entity.

Perceptual Filtering

In order for our memory not to be blocked by too much information, there must exist some selection of what memories are committed to be stored out of the ocean of sounds, colors and sensual perceptions which surround us daily. It seems likely that in an agent with eidetic memory, only part of his perceptual filtering process is active. The criteria by which information is filtered seem largely learned (Rose, 105). These criteria are themselves subject to change. This is part of the reason why humans are adaptable to very different environments. An example of such adjustment could be an individual's culture shock and subsequent adaptation to his/her new environment when living abroad in a country with very different customs.

Thus, remembering becomes an "active process," in which we sometimes very consciously choose the criteria of remembrance. Such self-construction of information-filtering criteria might account for a student's "subject-drawer" memory, by which it can be difficult for her/him to recall a fact about biology because she/he is currently sitting in an art class.

As Rose seems to suggest, perceptual filtering is the exclusion of information from being remembered. It seems to be a mechanism by which we have to forget in order to be able to survive in this world. This, of course, goes directly against Locke's account of memory. Eidetic memory is not the norm. We are designed, so to speak, to forget particulars of a present experience. For example, imagine if one were to ask the reader of this paper to recite, word by word, the sentences I wrote as my interpretation of Locke's quotation in the beginning of this essay. I would be very honored if the reader would easily reiterate every one of my statements. Yet, it is very likely that only a less detailed impression remained. Many words became forgotten particulars. Some of the words have not made it into a longer-term memory. In the same way, many of the smells, sights, and sounds of every moment of one's existence do not seem to be retrievable from memory. (For example, can one remember what one smelled at one's sixth birthday, or describe every pattern of the wrapping paper of one's Christmas presents of last year?) Some information about the present situation is not even committed to longer-term memory at all. If this is accepted from the outset, a retrieval of a "same consciousness of an event past" is made impossible because particulars of the past situation were not stored in memory.

Furthermore, the information that is stored as memory depends on the criteria of one's perceptual filtering. However, these criteria are subject to change themselves. Thus, not only is what is recalled in constant flux (because of mood/SWB dependency), but so is the information which is committed to memory.

That means that in very similar situations, at two instances (t1 and t2), the type of information remembered about event X may be different (due to SWB/MED dependency). Also, the information about very similar situations X1 and X2 (sitting in my favorite armchair tonight and tomorrow) that are committed to memory will probably be different, since between t1 and t2, one's filtering criteria might have (and are, in fact, very likely to have) changed.

A special, sad case of "forgetfulness" is Alzheimer's disease. The brain shrinks and thereby alters the structure of the neurons in the brain. This goes along with a constant loss of declarative memory (the memory that we can consciously recall). (The episodic memory [subgroup of declarative memories] especially suffers in this disease.)

The change in memory that concurs with the alterations in the neuronal structure of the brain seems to suggest that our memories are an encoding into the neuronal structure of the brain. This statement, based on empirical evidence about the progression of the physical changes within the brain as well as the loss of memory, is a mere induction from a number of cases. One could invent fantastic counter-theories, amongst which we could hardly decide using deductive tools. For example, a Cartesian might claim that certain changes in the brain will happen in parallel with changes in one's mysterious Cartesian ego void, where one's memories were placed. Yet, as an empirical fact, the claim that memory is solely encoded in the brain seems the most probable and supported truth.[14]

An interesting discovery was that patients suffering from Alzheimer's disease, when losing their episodic memory (the memory of personal experiences), can nevertheless retain their procedural (skill, habitual) memories. This again seems to be even further evidence for the distinction between descriptive and procedural memory. We are able to lose the conscious memory of events past, but we can retain skills that we have learned (without remembering, for example, the act of learning them).[15] Our memory is not unified, which principally makes it easy to allow for the occurrence of one aspect of memory and not another, both of which were gained from the same past experience. Let us look at another disease of the brain, something I wish to call ubiquitous Alzheimer's.

---

[14] It would now still be up to a Cartesian to show just how a thing out of space and time can influence an object in space and time, and furthermore, why the mysterious void happens to react in parallel with neuronal changes in the brain.

[15] In one case, an English piano player could still play some pieces on the piano, even though he had lost all of his declarative memory (memories about his piano lessons, performances, his friends, travels, etc.).

Ubiquitous Alzheimer's

When aging, many people start to lose memories. This seems due to the loss of some neuronal structures in the brain (Rose, 112). But exactly this fact led many doctors and scientists to the belief that there is such a thing as an "age-associated memory loss," an age disease. The neuronal changes that take place as a result of Alzheimer's are just more drastic. In fact, the National Institute of Mental Health "...has somewhat reluctantly come to the conclusion that the disease really exists...." Whether or not this is true, what is certain is that with age, mental processes change (Rose, 113).

It is important to notice that aging seems accompanied by a change in memory capabilities and an overall change of mental processes. This means that at a late stage of life, memory does not seem to be a static entity.

In summary, forgetfulness means that some information is lost without immediate retrieval. In the case of perceptual filtering and Alzheimer's disease, information is lost without any possibility of retrieval. The way in which one remembers, how and what one selects to remember, does not remain the same during one's life. The type of memory--the "color of memory," so to speak--is tied not only to mood (MED), SWB, repressed memories and unconscious mental states, but also to one's stage in life and one's perceptual filtering criteria. All these factors influence the character of one's memories. Upon reviewing such evidence, it seems difficult to maintain that memory is a static entity. Perceptional filtering is one of the factors that clearly seem to disprove even the theoretical possibility of a static sameness of memory. A personal identity demanding a static memory as a premise for its validity seems to be in difficulty.

Some biological considerations

It seems evident that biochemical changes take place when something is remembered (Klivington, 172). Synaptic changes take place in the brain when a thing is learned. We cannot yet exactly pinpoint how exactly memory information is encoded. However, research on children growing up in differing language communities has shown that our brain affords great plasticity in learning processes. For example:

> ...For most non-Japanese the sounds perceived by the left hemisphere (the side for speech) are actually linked to syllables containing consonants, while vowels, the noises of our environment, and the sounds that express an emotion...are perceived by the right hemisphere for non-Japanese, who use many more consonants in their speech, and by the left hemisphere for Japanese, who employ many vowels.... It was proven that these differences in brain patterns are caused by the environment, not by racial difference... (Klivington, 55).

Such evidence indicates that learning seems to directly manipulate the biochemical makeup of the brain.

With experiences and subsequent memories, biochemical changes take place.[16] It follows from this that the brain as a whole is far from being a static entity. This means that the brain itself ("the hardware") can also not be the centerpiece of an identity relationship (the hope of some physicalist theory), because the guarantee for a biochemical consistency over time seems not to be given.

A theory about memory

The categories I have used to explore knowledge about memory are certainly not exhausted, but I think strong evidence has amassed which can lead to a probable theory about memory. It is now time to explicitly connect the preceding empirical discoveries with my claims about memory. Let me restate the four points of a theory of memory introduced in the beginning of the paper, and let me give some reasons how they might follow from the empirical data I have surveyed so far.

(1) Memories are encoded in the brain.

The case of Alzheimer's disease, as well as the learning placidity of the brain, indicate that memories are encoded in the brain. In the former case, concurrent to observable brain degeneration occurs a loss of memory. In the latter case, synaptic changes are observable as a result of learning processes (Myers, 268). For example, the case of differing hemispherical dominance in Japanese and European brains (Klivington, 55) seems to indicate a dependency between exposure to, as well as retention of, certain information and biochemical changes in the brain. Such empirical evidence allows for a well-grounded induction, which would make claim (1) a very probable one.

2) Memory is constantly "mutating" on two plains--*vertical* (there are immediate factors that influence the memory recall), and *horizontal* (there is a long-term alteration of memory over time).

The dependency of mood and memory is one such indicator for the immediate mutability of memory (vertical). In horizontal memory, the change of memory type from eidetic to adult-like, the SWB-memory dependency, the changes of memory capabilities in old age, and finally, the alteration of perceptual filtering criteria are all factors which have a long-term impact upon memory.

(3) There is no "unique, original memory."

In the passage I cited in the outset of my paper, Locke seems to have contended that there is a stable, unique memory. Forgetting means only to forget parts or all of a unique, original memory. However, if one accepts claim (1), the endurance of an

---

16 Another indication for such claim is that "memories begin as impulses whizzing through brain circuits but then somehow consolidate into permanent neural changes" (Myers, 268).

original memory might not be possible: The persistence of a same neuronal basis for a memory seems doubtful, since the biochemical make-up of the brain is changing continually. If one accepts claim (1) and tries to reject claim (3), it must be shown that the biochemical encoding of a memory remains stable (while around this "pond" of stability, thousands of biochemical reactions take place).

However, if one accepts claim (2) and at the same time contends that the neuronal level remains somewhat constant, we could still never pin-point to an original memory, because the memory is altered vertically (due to mood [MED], physical well-being, etc.) and horizontally (long-term changes of perception criteria, SWB, aging, etc.). Now, if one were to reject claims (1) and (2), it must be asked just what explanation can be given for the empirical data I cited in favor of claims (1) and (2). In such a case, it is necessary to find an alternative storage medium for memory, which seems to lead back to a Cartesian void or "soul stuff." This poses the question of just how such wonderful "soul substance" is connected to an individual--in particular, how it is connected to this individual's brain. For example, if claim (1) is rejected, how can the change in memory be explained that concurs with alterations in the brain (as, for instance, during Alzheimer's disease)? Claim (2), on the other hand, can only be rejected if the empirical data I cited in its favor can be discounted.

I do not suggest that an agreement to claims (1) and (2) is necessary. However, their rejection might lead back to a version of Cartesianism with all its problems; as well, it might lead to a denial of the importance of a host of empirical evidence supporting claims (1) and (2).

If, however, claims (1) and (2) are granted, claim (3) seems readily to follow: biochemical alterations in the brain might prevent an unchanged storage of memory [extending claim (1)], just as vertical and horizontal changes of memory prevent one from pin-pointing a unique, original memory. This, in turn, means that there does not seem to be a "unique, original" memory.

(4) It is not possible to recall a memory in exactly the same way twice.

If claims (2) and (3) are accepted, then (4) must follow: If memory is changing vertically as well as horizontally, and if, in fact, we cannot even point to an original memory to which we compare other memories in order to establish their sameness to each other, then we cannot recall a memory in exactly the same way twice.

It is now of interest to see some of the consequences that the acceptance of the above claims (1)-(4) will have on a theory of personal identity based on memory.

According to claims (2)-(4), a pure Lockean notion of memory as being conscious as the "same self now and then" is rejected outright. It seems ruled out that we can remember an event as if it were when we experienced it [claims (2) and (4)].

An account of personal identity that is established on a static sameness of the re-experience of an event has to fail [(2) and (4)]. As I pointed out, there are many facts known about memory that do not coincide with such a claim: to name but one, perceptual filtering. If we accept that a theory of personal identity cannot be established on a static sameness of memory, we can only hope to retain some of the Lockean account of personal identity and try to fix it.

Parfit tried to amend the Lockean account of personal identity by introducing the Quasi-Memory. By doing so, he tried (with questionable success) to escape

Butler's critique of circularity.[17] Parfit thus stipulates that in order to draw the R-Relation, "...I seem to remember having an experience...someone did have this experience..." (Parfit, 220). What interests me here is less the someone (the circularity argument), but more, the *memory of having an experience.* Claim (4) shows that there does not seem to be any identical memory. In his *Venetian Memories* experiment, Parfit seems to presuppose just this: "...She ought to conclude that she has an *accurate* quasi-memory of how this flash of lightning looked to Paul..." (Parfit, 220). But as the cited samples from science and psychology show, there is no *accurate* memory at two points in time.

So what does it mean to remember having had an experience, let us say of event X? If we can never remember in exactly the same way, recalling X will lead to different memories. Thus, recall of event X at t1 is not identical to recall of event X at t2.

Of course, one might try to escape by assigning "crucial properties" to some X event, which thus is somehow supposed to make it the same X over time. But here claim (3) of the memory theory comes into play: How are we to decide on these crucial properties if we can never "fully, originally" remember event X? It seems much more likely that many vertical and horizontal influences do allow content properties of a recall of X to fluctuate. Bits of information that we carve out as more crucial than others are selected on the grounds of a changing grid. If one wants to derive personal identity from memory, one has to be more specific than assuming *the* memory of X by S at time t1 and assuming *the* memory of event X by S at time t2.

In fact, there does not exist a memory of an event X as a unique, self-contained entity. As the examples from science and psychology suggest, there are no clear-cut, untainted information marbles of a past event which simply roll from time to time into one's awareness. The collection of all episodic memories is not a filled marble sack.

Terminology such as "episodic memory" and "memory X, where X was such and such event" are somewhat misleading. These words falsely suggest that there exist stable self-contained entities of information over time that can be assigned to one's content of awareness at an earlier time. The content of one's momentary awareness is the sum of everything one senses, feels and thinks when, for example, reading the little word "this" right now. Yet, as was shown (as a result of perceptual filtering, for example), the remembrance of the event of reading the above little word "this" will only leave partial traces in one's memory, if any at all, as I hope.

Most defenders of the existence of personal identity, however, would allow for a curtailing of available information of the past and still hold it possible to draw an identity relationship based on unchanging bits of information. Remembering something is to commit some information (perceptual or descriptive) into a vast interconnected system of previously added information. The content of this information undergoes changes. This makes it necessary for personal identity theory defenders to reduce their premises to *some* information of the past time t1. But this

---

17 Butler argued that we cannot base personal identity on memory because every memory already presupposes a self, an "I," which we wanted to prove by it in the first place.

information does change also. For example, yesterday I saw the television show *Fawlty Towers* (time t1) and was extremely amused. Today (time t2) I remember only some of the scenes, but know that two friends visited and we watched the show together. Two years from now (time t3) I might only recall that sometime around April I watched something amusing, and cannot remember the friends' visit at all. The content of both memories (at t2 and t3) is quite dissimilar. Now, a defender of Personal Identity might say that it is enough that I know something happened. Yet it is interesting to note at this point that from the vast "Lockean continent" of "being conscious of one's past as it were present," almost everything was carved away. A defender of personal identity must flee onto the little island of affirming that we remember that at least *something* happened that is not a fiction.[18] But even this soil will become very watery, as I demonstrate.

We make an affirmation at time t2 of the truth of event X at time t1. For instance, three years after the afternoon with my friend: "Oh, I remember I watched Monty Python's *Life of Brian* three years ago...." In fact, we do often tell ourselves tales about ourselves[19] and turn them into truth affirmation, memories *that such and such*. If we did not have people around us, photographs, diaries, records of some kind to tell us tales about ourselves, we probably would end our lives with an even greater set of self-constructed fictions about ourselves. Thus, even if one tries to base an identity relationship on something as meager as a single affirmation of one crucial point of event X, one may be in vain.

In connection with claims (2) and (3), it might be noticed that there is no guarantee that a crucial patch of the properties of event X remains unchanged within our memories. To make an affirmation about event X might actually mean to affirm something that was not a part of the content of the momentary awareness at instance X. That means that such a simple affirmation of a content property of event X is not necessarily giving us the desired point of identity over time, which an account of personality identity over time should do. Also, a simple existence affirmation of event X will not help, since many individuals affirm the existence of an event X without the event X actually having taken place. For example, for a time I was convinced that I had spent a week in the mountains when I was 8 years of age; however, by photographs and by my own diary of that time, it was "proven" to me that I spent the whole of the summer at my house. Thus, the memory I had now of that summer affirmed an event which the memory which I had in the fall following the disputed summer could scarcely have affirmed. Also, the content of experience of the disputed summer was different, affirming a Y (staying home) from the content of memory until I opened the diary, affirming X (spending the week in the mountains).

---

[18] The word "fictitious" here refers not only to fictions about a real-life event, but also to fictions about previous memory experiences.

[19] In a study about the time estimates of past events in their life, many subjects tried to reconstruct the time endurance by available common-sense information (e.g., how long did I usually go on vacation), but this often led to wrong estimates, thus suggesting that affirmations *that...*, are bound to change as well (Burt).

Therefore, a simple existence affirmation--say, of event X in the past--is not enough to ascertain an identity relationship with a person at that time in the past, because event X might not have occurred, thus leaving no truly existent agent performing X with which an agent now could be rendered identical. Yet one might say that a remembrance of an X,[20] no matter if fictitious or not, at two points in time shall be enough to produce an identity relationship. But here again claims (2)-(4) come into play. How can we construct an identity relationship between a remembrance X1 and remembrance X2 at time t1 and t2, if we cannot remember in the same way twice and have no original, unique memory to form our basis of the identity relationship?

Let me be more explicit. If someone recalls that *such and such* happened as a part of event X twenty years ago (which it in fact did), can this be enough to assert that we should call the person at the event twenty years ago and the person now the same? I think it would be very difficult to do so: If, one day, one looked at a painting of a glaring red apple, yet twenty years later one only remembers the fact of having seen an object which was glaringly red, nothing much is said. In such a case, the affirmation of having seen something red would not distinguish between any of the near infinite number of possible situations that would fit this affirmation. I would affirm that fire trucks, Californian sunsets, and glaring red apples, etc., could have entered my visual field at the supposed event X twenty years ago. It seems that we need an affirmation (or set of affirmations) about event X, which undoubtedly ascertains a link into the past.

Not just any seemingly recalled property of event X is sufficient enough to support at least a meager connection into one's past. But here we again encounter the difficulties of finding just what these crucial affirmations should be. As already mentioned, claims (2) and (3) deny the certainty of a time-enduring, unchanging patch of content properties of event X in memory. Giving a set of affirmations and calling it crucial will not circumvent the possibility that some, or even all, of the set's affirmations are fictitious. It seems increasingly difficult to construct an identity link into the past out of one's first-person perspective.

But let us assume, for a moment, that such a link into the past is somehow possible. We have connected person A at event X at t1 with person B living at t2 and call them identical. Why should we then presuppose that everything in between the time of my recall t2 and t1 belongs to the same person? One cannot somehow escape by altering the identity premise for the remaining t2-t1 years and say that the person had to be tied to some living organism, and that I happen to be the continuation of this organism now, since I recalled X.[21]

To construct an account of personal identity based on memory, we might have two ways of going. One way is to posit a dense network of links between short periods of time, giving more certainty to the possibility of picking non-fictitious affirmations. Such a network would have to take into account the mutability of

---

[20] I want to remind once again that event X itself can be a memory experience.

[21] In *On Identity and Diversity,* Locke actually showed that we must not assume a bodily continuum for a person. Both are not linked together of necessity.

memory, especially the fact that linking affirmations are themselves prone to change, are placed at different times, places, circumstances, and are influenced by horizontal and vertical contingencies. The second way would be to come to the conclusion that the construction of a personal identity based on memory alone would involve insurmountable difficulties. Memory gives us the strongest tool in building a theory of personal identity, but we might be in need of something more. This "something more" might be "third-person accounts," which increase the validity of the set of crucial affirmations which is made about event X. Such objects could be diary entries, photos, tales of other persons.

On the whole, I am not suggesting that we cannot make any statements about ourselves. On the contrary, we are able to construct elaborate episodes about ourselves in almost all moments of our life. What I maintain, however, is that these episodes are far from being immutable. A theory of personal identity based on memory will be difficult to construct. It will have to take into account the severe changes that take place with the material that the connections into the past are made out of: mutating information and non-self-contained memories. A personal identity theory based on remembrance will not be able to ignore the continually changing "color of memory."

Bibliography

Borges, Jorge Luis, "Funes the Memorious," in *A Personal Anthology*, ed. by Anthony Kerrigan. New York: Grove Press, 1967.

Burt, Christopher D.B., "Reconstruction of the Duration of Autobiographical Events," *Memory and Cognition* 20, 2 (1992), 124-32.

Butler, Joseph, "Of Personal Identity," in *Personal Identity,* ed. by John Perry. Berkeley: Univ. of California Press, 1975.

Kandel, Minouche and Eric, "Flights of Memory," *Discover Magazine* 15, 4 (May 1994), 32-38.

Klivington, Kenneth, *The Science of Mind.* Cambridge, MA: MIT Press, 1989.

Myers, David G., *Psychology.* New York: Worth Publishers, 1992.

Parfit, Derek, *Reasons and Persons.* Oxford: Clarendon Press, 1984, Chapters 10-12.

Perry, John, ed., *Personal Identity.* Berkeley: Univ. of California Press, 1975.

Rose, Steven P.R., *The Making of Memory.* New York: Anchor Books Doubleday, 1993.

Schacter, Daniel, "Implicit Memory," *Journal of Experimental Psychology* 13 (1987), 501-18.

Seidlitz, Larry, and Ed Diener, "Memory for Positive Versus Negative Life Events," *Journal of Personality and Social Psychology* 64, 4 (1993), 654-64.

Squire, Larry, "Memory and Brain Systems," invited address to the American Psychological Society convention (1990).

Stevenson, Leslie, *Seven Theories about Human Nature.* Oxford: Oxford Univ. Press, 1987.

Taylor, Levi, "Relationship Between Affect and Memory," *Journal of Personality and Social Psychology* 62, 5 (May 1992).

Woodward, Kenneth L., "Was It Real or Memories?", *Newsweek,* March 14, 1994.