CHAPTER 11

Memory

Kourken Michaelian
Lecturer, Department of Philosophy
University of Otago, Dunedin, New Zealand

The past tells us who we've become. —Hauser, in *Total Recall*

The past is a construct of the mind. —Matthias, in *Total Recall*

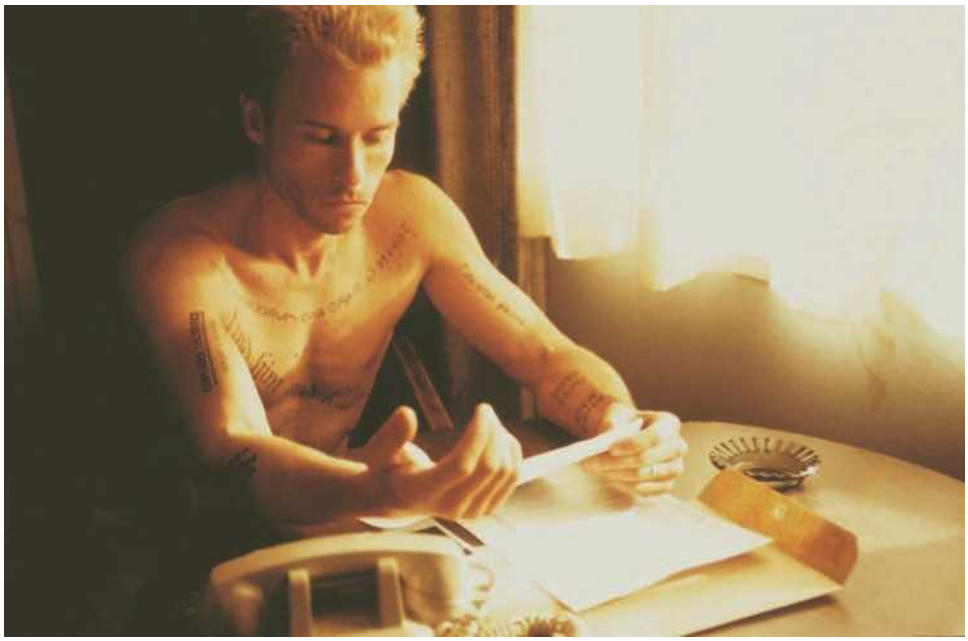
Memory can change the shape of a room or the color of a car. It's an interpretation, not a record. Memories can be changed or distorted and they're irrelevant if you have the facts. —Leonard, in *Memento*

Memory is one of our most fundamental cognitive capacities. Therefore, it should come as no surprise that it is also one of the oldest topics in philosophy. Linked to the metaphysical doctrine of the forms, it figured centrally in Plato's (427–347 BCE) theory of knowledge. And Plato's student, Aristotle (384–322 BCE), sought to distinguish memory from sensation and expectation by arguing that memory—and only memory—is about the past.

Although philosophers continue to discuss the relationship between memory and pastness, that question will remain in the background in this chapter. The bulk of the chapter will be devoted to a pair of interrelated questions, selected to be representative of current research on memory in the philosophy of mind. Our first question concerns the nature of *remembering*: What is remembering? In other words: What conditions need to be met before we may say that someone really remembers something? Our second question concerns the nature of *memories*: What are memories? In particular: Are the external memories stored, for example, in one's computer really memories in the same sense as the memories stored in one's brain?

The first of these questions is quite traditional, and philosophers historically have defended a range of different answers to it. We will review these answers, as well as a new answer that only recently has been proposed. The second question is more novel. Although debates over competing metaphors for memories go as far back as Plato and Aristotle, the question has taken on new life under the impact of recent controversies over the extended mind hypothesis. As we will see, our answers to these two questions interact. More rigid conceptions of remembering make it harder to count external memories as genuine memories, whereas more flexible conceptions make it easier to do so. By the same token, taking the notion of external memory seriously pushes us in the

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210



***Guy Pearce as Leonard in the 2000 film* Memento.** AF ARCHIVE/ALAMY STOCK PHOTO.

direction of a flexible conception of remembering as constructing, imagining, or simulating the past.

A 2000 film vividly illustrates the interaction between external memory and reconstructive remembering. In *Memento*, Leonard, the protagonist, suffers from severe anterograde amnesia (an inability to form new memories). He therefore relies on external records to reconstruct the events that have taken place since the accident that caused his amnesia. Leonard faces real difficulties because of his impairment: in an important sense, external memory is a poor substitute for internal memory. And yet, when he reconstructs past events on the basis of fallible external records, we have the sense that he really is remembering: after all, even unimpaired subjects often depend on external records to guide their remembering. Normally, we are capable of remembering without relying on external memory stores, but, even when we do so, our remembering manifests the reconstructive character illustrated by Leonard's struggles to remember. In a real sense, even internal memories provide only fallible clues to what happened in the past.

A second film illustrates the challenging theoretical questions raised by the fallibility of reconstructive remembering. In *Total Recall* (2012), the protagonist discovers that a set of memories that he was about to have artificially implanted correspond to his past as a secret agent more closely than do his apparently real memories of life as an ordinary worker, which themselves turn out to have been artificially implanted. Suppose that he were to have had the first set of memories implanted. When he later retrieved them, would he really remember being a secret agent? Intuitively, we want to say that he would not. As we will see, however, the process that would unfold were he to retrieve the implanted memories bears sufficient similarity to the process that unfolds when normal subjects retrieve their memories that a case can be made for saying that he would.

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

NARROWING THINGS DOWN

Before going any further, we need to narrow things down somewhat. Memory can refer to a broad range of different capacities, actions, and states of affairs. In line with philosophical tradition, our focus will be on one particular kind of memory, namely, episodic memory, but we need to distinguish episodic memory from other kinds of memory. Fortunately, philosophers and psychologists broadly agree on a taxonomy of kinds of memory.

NONDECLARATIVE VERSUS DECLARATIVE MEMORY

Adopting the vocabulary used in psychology, the basic division is between nondeclarative and declarative memory.

Nondeclarative memory, which philosophers often have referred to as habit memory, includes the kind of memory that is at work when one remembers how to do something. Remembering, in this sense, arguably does not require that one be able to make explicit or "declare" the content of what one remembers. I retain my ability to ride a bicycle, for example, even when I am not actually riding a bicycle, and in this sense, I can be said to remember how to ride a bicycle. When I get on a bicycle and ride it, my activity is in part the result of past experiences of bicycle riding, and in this sense, too, it is a form of nondeclarative memory. Note that this gives us two distinct senses—dispositional and occurrent—in which bicycle riding can involve memory. We will return to this distinction in the following paragraphs.

Although some recent controversy has arisen over this point, the standard view in philosophy is that knowledge of how to do something cannot be reduced to explicit, factual knowledge. Some forms of nondeclarative memory may be open to the influence of declarative memory, but nondeclarative memory also includes the simple forms of conditioning that you might have studied in a psychology class. Think, for example, of Ivan Pavlov's (1849–1936) famous dogs. The distinction between nondeclarative and declarative memory thus aligns roughly with the distinction between knowing how and knowing that. Knowing *that* something is the case is a matter of being related to a fact or proposition, and hence, when one knows that something is the case, one normally can articulate what one knows. Knowing *how* to do something, however, is a matter of having the skill or ability required to perform a certain kind of action, and hence, when one knows how to do something, one often will not be able to articulate what one knows. Similarly, one need not be able to make explicit what one nondeclaratively remembers, whereas one can, at least in principle, make explicit what one declaratively remembers (one can *declare* it). Thus, it makes sense to speak of retrieving information or content from declarative memory but not from nondeclarative memory.

SEMANTIC VERSUS EPISODIC MEMORY

The remainder of this chapter sets aside nondeclarative memory to focus on declarative memory, but we need to narrow things down slightly further before reaching our target.

Within declarative memory, the basic division is between semantic memory, which philosophers often have referred to as propositional memory, and episodic memory, which philosophers have often referred to as recollective memory. Semantic memory is the kind of memory that is at work when one remembers a proposition or fact. I continue to know that Paris is the capital of France, for example, even when I am not actively thinking of that fact, and in this sense, I can be said to remember that Paris is the capital of France. When I now think to myself that Paris is the capital of France, I am drawing on information acquired in past learning

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

experiences, and my thought is thus an act of memory. Again, note the two distinct senses (dispositional and occurrent) in which one can be said to *semantically* remember.

Episodic memory is the kind of memory that is at work when one remembers the events of one's personal past. I continue to remember my first trip to Dunedin (where I now live) even when I am not actively thinking of that event, and in this sense, I can be said to remember the trip. When I now think of the trip, I am drawing on information acquired through my experience of the trip, and my thought is thus an act of memory. Again, just as there are two senses in which one can be said to nondeclaratively remember or to semantically remember, there are two distinct senses (dispositional and occurrent) in which one can be said to *episodically* remember.

What episodic memory, semantic memory, and nondeclarative memory have in common is that they depend on past experience. Another way of putting this is to say that each form of memory is the outcome of a corresponding form of learning. But the relation to past experience is importantly different in each case. In the case of nondeclarative memory, the relation to past experience is implicit, because nondeclarative remembering does not require the retrieval of content that might refer to experiences through which memories were formed. In the case of semantic memory, the relation to past experience remains implicit: although semantic remembering does require the retrieval of content, the content in question need not refer to the experience through which the memory was formed. When I remember that Paris is the capital of France, my memory need not, and normally will not, make any reference to the event or events in which I learned that Paris is the capital of France. In contrast, when I remember my first trip to Dunedin, my memory necessarily points back to the event itself. Thus, only in the case of episodic memory is there an explicit relation to past experience.

Interestingly, this means that it is not redundant to say that one episodically remembers a given episode, whereas it is redundant to say that one semantically remembers a given proposition. I can remember that the French Revolution began in 1789, and I can remember that I traveled to Dunedin in 2014. Regardless of whether I remember a proposition that has nothing to do with me, as in the former case, or a proposition about an event from my own past, as in the latter case, the same, semantic form of memory is at work. In addition to semantically remembering that I first traveled to Dunedin in 2014, however, I can episodically remember the event of traveling there. Conversely, I cannot, no matter how much information I bring to mind, episodically remember the storming of the Bastille in 1789. Episodic memory thus requires something more than semantic memory. One possibility is that even if semantic memory can provide information or content that refers to one's past, episodic memory alone provides a distinctive sort of past-oriented episodic content. Another possibility is that no fundamental distinction exists between episodic and semantic content; instead, episodic memory alone provides one with a distinctive feeling of reliving past events. Philosophers have tended to favor the former possibility, often building an explicit reference to the past into their analyses of episodic memory content. Psychologists tend to favor the latter, in part on the basis of clinical cases in which a loss of past-oriented episodic phenomenology results in impaired episodic memory, even in patients who retain the ability to represent events from their pasts.

EPISODIC VERSUS AUTOBIOGRAPHICAL MEMORY

Episodic memory has always been the form of memory with which philosophers are centrally concerned; indeed, many have argued that, strictly speaking, only episodic memory is really

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

memory. Episodic memory will be our target in this chapter, but, before turning to theories of episodic remembering, we need to distinguish episodic memory from autobiographical memory.

Episodic memory is devoted to specific events or episodes from the personal past. (This raises an important question: What are episodes, in the relevant sense? Surprisingly, there is not a lot of work on this question.) Some controversy exists over this point, but, in the view of many psychologists and neuroscientists, episodic memory depends on a single dedicated system at the level of the brain. Autobiographical memory, in contrast, is memory for the personal past as a whole, including not only specific events (e.g., my first trip to Dunedin) but also repeated events (e.g., my trips to Dunedin in general) and whole periods of one's life (e.g., the period during which I lived in Dunedin). Unlike episodic memory, autobiographical memory does not depend on a single brain system. Instead, it is a complex capacity depending on multiple interacting systems. On the influential self-memory system framework, for example, autobiographical memory involves semantic knowledge of the overall structure of one's life, knowledge of specific periods within that structure, and phenomenological records. Only the latter, which represent specific events, correspond to episodic memory.

In everyday contexts, we care about episodic memory in part because of the way it contributes to autobiographical memory, which gives context and meaning to individual episodes. When the *Total Recall* character Hauser says that "the past tells us who we've become," he has autobiographical memory in mind. Nevertheless, we will focus on theories of episodic memory, rather than autobiographical memory more broadly.

WHAT IS REMEMBERING?

Let us turn, then, to our first question. What is episodic remembering? What is it for someone to remember an event?

Just as there are two senses in which one can be said to remember how to do something and two senses in which one can be said to remember a proposition, there are two senses in which one can be said to remember an event. We can refer to these senses as occurrent and dispositional. The theories on which we will focus are theories of occurrent episodic remembering. To occurrently remember an event, it is necessary that one consciously entertains a representation of that event. In principle, we also could inquire into dispositional episodic remembering. To dispositionally remember an event, it is necessary to have the capacity to form a conscious representation of the event, but it is not necessary that one actually does so. This means that one cannot occurrently remember an event unless one previously dispositionally remembered it, but one can dispositionally remember an event even if one never occurrently remembers it.

To illustrate the distinction between dispositional and occurrent remembering, return to the example of my first trip to Dunedin. When we say that I continue to remember the trip even when I am not actively thinking of it, we mean that I retain the capacity to remember it; what we have in mind is dispositional memory. In contrast, when we say that I am now remembering the trip, what we have in mind is occurrent memory. If I occurrently remember the trip, then I must have dispositionally remembered it. But I might dispositionally remember the trip without ever occurrently remembering it. This might happen if I have not forgotten the trip, in the sense that I retain the capacity to think about it, but for some reason am prevented from actually thinking about it. Freudian (after Sigmund Freud, 1856–1939) repressed memories, if they exist, provide a more dramatic illustration of dispositional remembering without occurrent remembering.

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

THE EMPIRICIST THEORY

To occurrently remember an event from the personal past, it is necessary to think of the event. But thinking of an event from the personal past is not sufficient for remembering it. Suppose that I suffer from retrograde amnesia (an inability to retrieve previously formed memories, a sort of mirror image of the anterograde amnesia discussed previously) and, as a consequence, have entirely forgotten my first trip to Dunedin. Later, I imagine, on the basis of my knowledge of more recent trips, a trip to Dunedin that happens to match my first trip in every detail. Do I count as remembering my first trip to Dunedin? Intuitively, I do not, and I would not even if I were to describe what I am doing as trying to remember the trip. Remembering an event involves more than merely imagining it. When we ask what it is for someone to remember an event, we are asking, in part, what it is for him to remember an event rather than to imagine it.

The question of what objectively makes the difference between remembering and imagining needs to be distinguished from the question of what enables us to subjectively tell the difference between remembering and imagining. It is one thing to ask: what determines whether I am remembering or imagining? It is another to ask: how can I determine whether I am remembering or imagining? There may be criteria that distinguish between remembering and imagining but that the subject cannot himself use to distinguish between remembering and imagining. And there may be criteria that the subject can use to distinguish between remembering and imagining with some reliability but that are not sufficiently reliable to mark the objective distinction between remembering and imagining. The second of these questions, the question of how subjects themselves avoid confusion between remembering and imagining, belongs to epistemology. We focus on the first question, which belongs to the philosophy of mind.

David Hume's (1711–1776) *empiricist* theory of memory proposes two separate criteria for memory. According to the first criterion, remembering is distinguished from imagining by the greater degree of flexibility characteristic of the latter. The idea is that, in remembering, we are constrained by the shape of the original event, whereas, in imagining, we are free to construct a representation of an event as we see fit. Although it initially appears plausible, ultimately, this idea cannot work, for remembering is itself a highly flexible process. We will explore this point in more detail, but, to illustrate, consider the possibility of switching between field and observer perspectives in remembering. In *field* memory, one remembers a scene by seeing it from the perspective from which one originally experienced it. In *observer* memory, one remembers the scene from the perspective of a hypothetical observer, potentially seeing oneself in the scene. Observer memory is obviously genuine memory, but it involves considerable flexible reconstruction of events.

Hume's second criterion fares little better. According to this criterion, remembering is distinguished from imagining by the greater vivacity of the former. This idea can be interpreted in two ways, neither very promising. If, on the one hand, the representations produced by remembering are characterized by a greater level of detail than the representations produced by imagining, it cannot work. Representations produced by remembering may be more detailed on average than representations produced by imagining, but only on average. Choose an event that you remember only dimly. It should be easy to imagine a hypothetical event in more detail than you remember this half-forgotten scene. If, on the other hand, remembering imposes itself on us more forcefully than does imagining, the vivacity criterion, again, cannot work. Remembering sometimes takes a great deal of work. Again, recovering Freudian repressed memories may provide an illustration. And

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

imagining sometimes imposes itself on us quite forcefully, as when one is distracted by daydreams while attempting to focus on some mundane task.

THE EPISTEMIC THEORY

One natural alternative to the empiricist theory is the *epistemic* theory. According to the epistemic theory, memory is, roughly, retained knowledge.

The epistemic theory is usually treated as an account of semantic memory rather than an account of episodic memory. Taken as an account of semantic memory, the theory says that to remember a proposition is to know it and to have known it continuously since one initially acquired it. If we can agree on a suitable conception of episodic knowledge, however, the epistemic theory can be extended to provide an account of episodic memory in a straightforward manner. Taken as an account of episodic memory, the theory will say that to remember an episode is to have knowledge of the episode and to have had that knowledge continuously since one initially acquired it. Although we ultimately are concerned with episodic memory, it will simplify our discussion to focus on the standard, semantic version of the epistemic theory.

Several problems make it difficult to spell out a satisfactory version of the epistemic theory. The first problem is that it is difficult to formulate a version of the theory that is neither too strong (too demanding) nor too weak (too undemanding). Stronger versions of the theory tend to rule out the possibility of remembering something after having forgotten it. If I temporarily forget some information but later can retrieve it, I should, at that later time, count as remembering. Strong versions of the epistemic theory, however, imply that I do not. Weaker versions of the epistemic theory present the opposite problem and tend to rule out the possibility of genuinely relearning something. If I forget some information completely but later relearn it, I should not, at that later time, count as remembering. Weak versions of the theory, however, imply that I do.

A second problem for the epistemic theory concerns its inability to accommodate the possibility of remembering without knowing. There are three aspects to this problem. First, it is possible to remember without knowing in the sense of performing an act of retrieval that qualifies as remembering but for which the belief that is retrieved does not meet all the conditions for knowledge when it is retrieved. Knowledge implies both justification and truth. Justification is fallible. When one's justification for a given belief is overridden, we say that it is defeated; the information that overrides it is a defeater. So one might form a justified true belief that later becomes unjustified because one acquires information that defeats one's justification. In cases of this sort, a belief that initially qualifies as knowledge later, under the impact of a defeater, ceases to qualify as knowledge. For example, I might form the belief that the French Revolution occurred in 1789 by reading about it on a website that I believe to be reliable. In this case, I have a justified true belief that the French Revolution occurred in 1789. Later, someone might tell me that the website in question is unreliable. I might stubbornly continue to believe that the French Revolution occurred in 1789. Although it is in fact true that the French Revolution occurred in 1789, my justification for this belief has been defeated. This gets the epistemic theory into trouble. If I acquire defeaters for a belief while it is held in memory, the belief, which initially might have satisfied all the conditions for knowledge, will fail to do so when retrieved, because it fails to be justified. Thus, a belief that initially qualified as knowledge fails to do so when retrieved. According to the epistemic theory, because I do not count as knowing, I do not count as

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

remembering. Intuitively, people should be able to remember even if their justification has been undermined.

Second, it is possible to remember without knowing in the sense of performing an act of retrieval that qualifies as remembering but for which the belief in question never satisfied the conditions for knowledge to begin with. Suppose that I form a belief that is true but not justified. For example, I might not know when the French Revolution occurred. Being superstitious, I decide to find out by reading tea leaves. Coincidentally, the tea leaves spell out "1789," and I form a true but unjustified belief about the date of the French Revolution. Time passes and I do not gain or lose any defeaters for my belief. Later, I retrieve the belief. According to the epistemic theory, because I do not count as knowing, I do not count as remembering. Intuitively, however, a person should be able to remember even if they never had justification.

Third, it is possible to remember without knowing in the sense of performing an act of retrieval that qualifies as remembering but for which the belief in question fails to satisfy the conditions for knowledge because it fails to be true. Because the epistemic theory defines memory in terms of knowledge, and because, as we have seen, knowledge implies truth, the theory says that memory implies truth. Theories of memory that say, in this manner, that one cannot remember a proposition unless the proposition in question is true are said to be *factive* (because they say that one can only remember facts). The claim that memory is factive has been challenged, but many philosophers have found it to be plausible. Now, it is true that in ordinary conversational contexts, one person can legitimately challenge another person's claim to remember by pointing out that what the other person claims to remember is false. But this fact about ordinary usage should be given little weight in a philosophical context, in which our aim is to characterize not the ordinary usage of the verb *to remember* but rather the nature of remembering itself. Suppose that I come to falsely believe that the French Revolution occurred in 1689. (For present purposes, it makes no difference whether my belief is justified or unjustified.) Suppose that I later retrieve this belief. According to the epistemic theory—like any theory that views memory as factive—this act of retrieval does not count as an act of remembering. But there need be no difference, at the cognitive level, between what happens when I retrieve a false belief and what happens when I retrieve a true belief. Thus, the fact that you might challenge my claim to remember by pointing out that my belief is false does not reveal whether or not I am remembering.

It is possible to make any of the other theories of remembering considered thus far into a factive theory simply by requiring that what the subject supposedly remembers be true before the subject can be said to really remember. But factivity poses a special problem for the epistemic theory. Because that theory defines memory in terms of knowledge, and because truth is essential to knowledge, any version of the epistemic theory along the lines of that considered thus far is subject to the problem of nonfactive remembering. Other theories of remembering, in contrast, can dodge the problem of nonfactive remembering simply by opting not to include the factivity requirement.

A final problem for the epistemic theory concerns its explanatory value. If a definition of memory as retained knowledge is to be at all informative, we need to know what it is to retain knowledge. The only apparent way to unpack the idea of retention of knowledge is in terms of a causal connection between an earlier belief and a later belief. Thus, the epistemic theory seems to collapse into the causal theory, to which to we turn next.

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

THE CAUSAL THEORY

Consider the example of my first trip to Dunedin. Suppose that, after experiencing the trip, I suffer an accident that, giving rise to severe retrograde amnesia, wipes all trace of the trip from my mind. Suppose that I later go to see a hypnotist. Lacking any knowledge of my trip to Dunedin, the hypnotist coincidentally hypnotizes me into believing that I made a trip to Dunedin that matches my actual trip in every detail. Intuitively, despite this detailed representation, I do not remember the trip. According to the causal theory of memory, the reason I do not remember it is that no causal connection exists between the representation and the event: what makes the difference between remembering and imagining is that remembering an event, unlike imagining it, requires a causal connection between the subject's current representation of the event and the subject's original experience of the event.

Not just any causal connection between the original experience and the current representation will do, however. Consider again the example of my first trip to Dunedin. Suppose that, after experiencing the trip, all trace of the event was wiped from my mind by an accident that gave rise to retrograde amnesia. Suppose, however, that, before suffering the accident, I told one of my colleagues about the trip in great detail. Suppose that, after the accident, my colleague told me about the trip, relating the same details that I told him. Suppose, finally, that I later completely forget that my colleague related information about the trip to me and now just recall the information itself. Intuitively, although a causal connection exists between my original experience of the trip and my current representation of it, I do not remember the trip.

To rule out this sort of inadequate causal connection, the causal theorist might say that remembering requires not just any causal connection between the current representation and the original experience but, specifically, a causal connection that goes via a memory trace. It is not enough, however, to say simply that remembering requires that the causal connection go via a memory trace. In the case just described, I have a memory trace of my trip to Dunedin up to the point at which I suffer the accident giving rise to my amnesia; it is this memory trace that allows me to tell my colleague about the trip. But this means that the causal connection between my original experience of the trip and my current representation of it initially go via a memory trace. So, to rule out this sort of inadequate causal connection, the causal theorist must say, more strictly, that remembering requires not just any causal connection between the current representation and the original experience, but, specifically, a causal connection that goes continuously via a memory trace. Setting aside the technical details, the idea is that the memory trace needs to exist all along—it must not be lost at some point along the causal chain—and it must contribute to the subject's act of recollection.

This raises an important question: What are memory traces? In principle, the notion of a memory trace is clear enough: a memory trace is a sort of record laid down by the subject's experience of the event. In practice, however, the nature of memory traces has been extremely controversial. In philosophy, much of this controversy has concerned the question of how the trace represents the original event. In psychology, much of the controversy has been about whether traces are discrete entities stored in specific locations in the brain. Although space does not permit going into these controversies in any detail here, we do need to discuss one intriguing possibility that, until recently, has not been taken seriously.

Most discussions of memory traces in philosophy and psychology have assumed that, whatever their nature, traces are internal records of events. But must they be internal? Could

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

memories instead be stored externally? When the causal theory was first introduced, its proponents refrained from assuming that traces have to be internal—discussing, for example, a hypothetical case of extraterrestrial beings whose memories are stored in detachable artificial devices. They argued that we have no reason to say that such beings would be incapable of genuinely remembering. Although most subsequent developments of the causal theory have assumed that traces are internal, recent discussions of external memory give us reason to revisit this question.

THE SIMULATION THEORY

In recent decades, the causal theory has been one of the most popular accounts of remembering in philosophy. But research conducted in psychology puts considerable pressure on the theory. This research is in the broad tradition of British psychologist Frederic Bartlett's (1886–1969) influential work on the constructive character of remembering. In his most famous experiment, Bartlett had his subjects read and later remember a folktale from an unfamiliar cultural context. He found that, as time went by, they tended to recall fewer details of the story. More interestingly, he found that subjects tended to modify the tale so that it better conformed to the narrative conventions of their own culture.

In line with this finding, subsequent research on constructive memory often has emphasized ways in which construction can give rise to inaccurate memories. Although construction no doubt sometimes gives rise to inaccurate memories, it does not necessarily do so. If the principles governing the process of construction are reliable, constructed memories might be accurate as often as—or more often than—they are inaccurate. Consider the phenomenon of boundary extension, in which subjects see part of a scene and subsequently remember more of the scene than they saw. For example, subjects might see a sheep with a field in the background and later remember more of the field than they actually saw. If the principles governing boundary extension are reliable, then this form of construction might produce accurate memories more often than not. For example, if the way in which subjects extend the representation of the field is accurate with respect to the field, their memories will be accurate, despite being constructed.

If it is possible, in this manner, for people to remember more of an event than they experienced, then we may need to give up the causal condition. A simulation theory of remembering emphasizes that the subject's current memory representations are always the product of a constructive process and does not require that these current representations be linked by an appropriate causal connection to the subject's original experience. According to the simulation theory, the very premise of our discussion so far is wrong. We have been assuming that the job of a theory of remembering is to distinguish remembering from imagining. But if remembering the past is a matter of simulating the past, there is no difference between remembering and imagining—to remember the past just is to imagine it. In an important sense, the past really is "a construct of the mind," as the character Matthias says in *Total Recall*.

We, however, may still distinguish remembering the past, viewed as a form of episodic imagination, from related forms of episodic imagination. Much of the support for the simulation theory of remembering comes from current research in psychology on mental time travel. Building on Bartlett's view of remembering as reconstruction, mental time travel research views remembering as one form of mental time travel among others. A wealth of evidence shows that the brain regions responsible for our ability to remember past events

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

overlap heavily with the brain regions responsible for our ability to imagine probable future events. Indeed, these regions also appear to be responsible for our ability to imagine alternatives to past events, alternatives to present events, and a whole range of improbable future events, suggesting that remembering is one product among others of a general episodic imagination system.

Philosophers are only now developing the simulation theory of episodic remembering, but an important challenge for the theory is to explain what distinguishes remembering, viewed as mental time travel into the past, from other forms of mental time travel into the past. One view is that it is the subject's intentions that do the work—you are remembering a past event, rather than imagining it in some other way, if you aim to form an accurate representation of the event. But remembering often is triggered involuntarily or automatically, and, in cases of involuntary memory, the subject has no intentions to which we might appeal. The simulation theorist owes us a workable alternative to this view.

WHAT ARE MEMORIES?

One interesting feature of the simulation theory is that it may require us to abandon intuitions about whether certain cases are cases of genuine remembering. Consider again the case in which I tell a colleague about my trip to Dunedin; lose my memory of the trip; am told about the trip by the colleague who learned about it from me; lose my memory of being told by the colleague; and, finally, "remember" the trip on the basis of the information transmitted by the colleague. Intuitively, in this case, I do not count as remembering the trip. If we assume that memory traces have to be internal to the subject—records of experienced events stored in the subject's own brain—the causal theory supports this intuitive judgment, because the causal connection between my current representation of the trip and my original experience of it does not go continuously via an internal memory trace but rather takes a detour via my colleague's memory. The simulation theory, unlike the causal theory, does not have this implication, because it acknowledges that subjects may draw on whatever information is available to them when constructing memories of past events.

As noted, however, it is open to causal theorists to acknowledge the existence of external memory traces. And if these theorists opt to do so, they may be in a position to allow, like the simulation theorist, that I genuinely remember my trip, despite my dependence on externally stored information. This brings us to our second set of questions. What are memories? Can memories be externally stored, or are they by definition internal to the subject?

THE CASE AGAINST EXTERNAL MEMORY

The case just described is one in which one human subject's ability to remember depends on information stored by another human subject, but we also need to consider cases in which a human subject's ability to remember depends on information stored by a technological resource, such as a notebook or a smartphone. Influential arguments for the extended mind hypothesis—the claim that mental states or cognitive processes might literally extend beyond the boundaries of the brain to include external resources—appeal to our intuitions about examples of both sorts. Our question in this chapter is about the notion of external memory, rather than about the extended mind hypothesis. That is, we are not concerned

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

with the general question of whether mental states or cognitive processes can involve external resources but with the specific question of whether memories can be stored externally. In principle, even if the extended mind hypothesis is right, we might still need to reject the notion of external memory: the external resources involved in these examples might be involved in mind or cognition without qualifying as memories, strictly speaking. But cases of supposed external memory have figured centrally in arguments for and against the extended mind hypothesis, so the questions are intertwined.

The most famous example of technological external memory is that of Otto, an imaginary Alzheimer's patient, who, to compensate for his failing memory, makes heavy use of a notebook that he constantly keeps with him. Whenever he receives some information he thinks he might need to know in the future, he writes it down in the notebook. Whenever he needs to know something, he looks it up in the notebook. Suppose that Otto wants to see an exhibition at a museum; he will consult his notebook for the address of the museum and then set off to find that address. Compare this to what happens when Inga, a subject with a normal memory, wants to see the exhibition; she will consult her memory for the address and then set off to find it. Extended mind theorists contend that the respective situations of Otto and Inga are strictly analogous: Otto's notebook plays the same role, with respect to his mind, as the relevant parts of Inga's brain play with respect to hers, so it should count as part of his mind, just as much as they should count as part of hers. And, they suggest, if the information stored in Inga's brain counts as a memory, so should the information stored in Otto's notebook. As long as everything else is the same, the mere location of the information is irrelevant.

Otto's case and other similar cases are used to support the *parity principle*, the claim that, if an external resource performs a task that would be counted as cognitive if it were performed by an internal resource, then the external resource is just as much engaged in cognition as the internal resource would be. The parity principle has been the target of numerous arguments by opponents of the extended mind hypothesis. But before considering these, let us review a case of social, rather than technological, external memory. Extended mind theorists argue that, for example, in an extremely interdependent married couple, someone else's beliefs might take the role played for Otto by his notebook and for Inga by the relevant parts of her brain; that is, some of your beliefs might be stored in someone else's brain. In such a case, the parity principle will again count the subject's cognitive processes as extending beyond his brain—in this case, into the brain of another person, instead of a technological artifact. Similarly, we might count the other person's beliefs as constituting some of the subject's own memories.

Both sorts of external memory—technological and social—have been subjected to attacks, but the effectiveness of particular attacks may depend on the kind of external memory in question. One sort of attack appeals to differences in the storage format of internal and external memory. This attack is most effective with respect to technological external memory. As discussed, part of the controversy over memory traces is about whether they are discrete, static entities laid down at specific locations in the brain. Without going into the details of this controversy, suffice it to say that its outcome was simple: they are not. Rather than being a matter of retrieving discrete representations stored at specific locations, remembering is a reconstructive process that depends on distributed and overlapping representations. Even when in storage, these representations are not static; because of their overlapping character, they constantly are being modified and updated under the impact of the subject's ongoing experience. Technological external memory is nothing like this.

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

Notebooks, smartphones, and the like are useful precisely because they store discrete, static traces. So the suggestion that a notebook can simply be swapped in as a replacement for a failing memory seems like a nonstarter. This point can be used to undermine parity-based arguments for the extended mind hypothesis. More important, for our purposes, it seems to undermine the idea of technological external memory.

Another sort of attack points not to differences in storage format but rather to differences in accessibility between internal and external memory. This attack is especially effective with respect to social external memory. An extremely interdependent couple might spend a great deal of time together, but one's access to the contents of another person's mind is never as reliable as one's access to the contents of one's own mind, even when the latter is limited by Alzheimer's disease or a similar impairment. One never endorses the memories of another as readily as one endorses one's own memories. Nor do the memories of another person, unlike one's own memories, result from one's previous endorsement of information. To be sure, couples do help each other to remember, prompting and correcting each other's recollections. But even if you are a member of a long-standing, closely interacting couple, the memories of the other person contribute to your own remembering precisely because they are different from, and complementary to, your own memories. So the suggestion that another person's mind can simply be swapped in as a replacement for a failing memory likewise seems like a nonstarter. Again, this point seems to undermine not only parity-based arguments for the extended mind hypothesis but also the idea of social external memory.

THE CASE FOR EXTERNAL MEMORY

These points need not, however, spell doom for the idea of external memory. Over time, extended mind theorists have shifted from parity-based arguments to complementarity-based arguments. Whereas parity-based arguments appeal to analogies between the roles played by internal and external resources in cognition, complementarity-based arguments appeal to disanalogies. In the same spirit, we can argue for external memory as a legitimate category despite the important disanalogies between the roles played by internal and external resources in remembering. The thought expressed by the character Leonard in *Memento* is mistaken. Memories—internal memories—clearly are not irrelevant as long as we have "the facts," that is, externally stored records. If they were irrelevant, memory impairments, such as Alzheimer's disease and amnesia, would not be the tragedies they are. (Leonard's own fate nicely illustrates this point, although, to avoid spoiling things for those who have not yet seen the film, it will not be described here.) The directly opposed thought, however, is likewise mistaken. Clearly, external records are equally not irrelevant as long as we have our internal memories.

As noted previously, forms of technological external memory such as notebooks and smartphones are useful precisely because they store information in a format that differs radically from that used by internal memory. If I want to remember my trip to Dunedin in vivid detail, I rely on my internal memory, which allows information originating in experiences of subsequent trips to the same destination, similar trips to other destinations, and general knowledge about air travel to bleed through. If I want to remember the exact time of my flight, in contrast, I would do best to look it up on my smartphone, where a copy of my itinerary is still stored. The two kinds of information may interact, each contributing something to the final product, a constructed representation of the past episode in which I first traveled to Dunedin. Remembering is often a process that depends on this

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

sort of interaction among multiple resources, internal and external, with each contributing something to the conscious memory representation.

This responds to the objection to technological external memory discussed earlier, which turned on differences in storage format. What of the objection to social external memory, which turns on differences in accessibility? A response in the spirit of complementarity is again available. Whereas parity-based arguments focus on cases in which the agent treats the information stored in another's mind just as he treats the information stored in his own mind, complementarity-based arguments highlight the multiple different ways in which we remember though conversation. As noted, it is often useful to remember with another precisely because their memories differ from your own. The ease with which our memories are modified by suggestions made by others suggests that we are designed not just to rely on those with whom we have long-standing relationships but also to flexibly update our memories under the influence of our current conversational partners in general, even if the relevant relationships are short lived. Even when we do not blindly trust information provided by others, such information may contribute in a variety of ways to our own remembering—for example, by providing cues that stimulate us to recall further information on our own.

OTHER QUESTIONS

The questions on which we have focused in this chapter are representative of current research, but a host of additional questions about memory are of concern to philosophers of mind. Three examples, picking up on the ideas discussed earlier, will illustrate the scope of research on memory in contemporary philosophy of mind. The work on different kinds of memory described in the section "Narrowing Things Down" raises a question about the relationships among these kinds of memory. Is memory fundamentally unitary, or is it a mistake to apply the same label to such disparate cognitive process? The different theories of remembering described in the section "What Is Remembering?" (and especially the simulation theory) raise a question about the accuracy of memory. In light of the thoroughly reconstructive character of remembering, retrieved memories are unlikely ever to correspond perfectly to past experience. What, then, are the appropriate standards of truth in memory? Finally, the debate over external memory, especially social external memory, reviewed in the section "What Are Memories?" raises a question about collective memory. Although we often remember in groups, does it ever make sense to talk, as we sometimes do, about groups themselves as remembering?

An important question that remained in the background is that of the relationship between memory and pastness noted by Aristotle. Many philosophers have favored an answer to this question in terms of memory content. The idea is that memory is about the past in the sense that the contents of memories explicitly point back to past events. But recent work on mental time travel supports an alternative answer, according to which memory is about the past in the sense that it involves a subjective feeling of pastness. Two additional important questions that we have not had time to discuss deserve to be noted. First, there is the question of the role of memory in determining personal identity. Personal identity is a matter of what makes a person count as the same person over time. One commonsense view on personal identity is that what determines personal identity is continuity of memory: you are the same person now as you were ten years ago because you remember the experiences of the person who existed ten years ago. Given the reconstructive

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

character of remembering, however, it is not clear that memory can provide the right sort of continuity to underwrite personal identity. Second, there is the question of the objects of memory. This is the question of whether what we are directly related to, when we remember, is the past episode we remember or rather the internal representation by means of which we remember it. Common sense speaks in favor of the idea that the objects of memory are past episodes, but, again, the reconstructive character suggests that our relation to past episodes may be quite indirect.

The works listed in the bibliography provide further reading on these and other questions about memory in the philosophy of mind.

Summary

This chapter discussed two main questions. First, what is remembering? We reviewed several traditional answers to this question—those given by the empiricist theory, the epistemic theory, and the causal theory—before concluding that a more unorthodox answer, that given by the simulation theory, looks promising. Second, what are memories? We reviewed arguments both for and against the view that memories can be externally stored, concluding that we may have good reason to think that they can be. As noted at the beginning of the chapter, our answers to these questions interact. We have seen that it is possible for causal theorists to endorse the idea of external memory, as long as they do not stipulate that memory traces are internal by definition. But an argument for external memory in the spirit of complementarity fits best with the simulation theory, which emphasizes the fact that remembering is always an imaginative process, a process of constructing a representation of a past event on the basis of whatever information is currently available, a process distinguished from other forms of imagination not by a special causal connection to the past event in question but simply by the fact that it aims to produce an accurate representation of that event. Just as it is possible for the causal theorist to endorse the idea of external memory, it also is possible for the simulation theorist to reject it. But the two ideas—simulational remembering and external memory—go naturally together.

Bibliography

Abbott, Barbara. "Linguistic Solutions to Philosophical Problems: The Case of Knowing How." *Philosophical Perspectives* 27, no. 1 (2013): 1–21. A review of the recent controversy over the relationship between knowing how and knowing that.

Anastasio, Thomas J., Kristen Ann Ehrenberger, Patrick Watson, and Wenyi Zhang. *Individual and Collective Memory Consolidation: Analogous Processes on Different Levels*. Cambridge, MA: MIT Press, 2012. Argues that collective memory is really memory.

Bartlett, Frederic C. *Remembering: A Study in Experimental and Social Psychology*. Cambridge: Cambridge University Press, 1932. The founding text of research on constructive memory.

Bernecker, Sven. "Memory Knowledge." In *The Routledge Companion to Epistemology*, edited by Sven Bernecker and Duncan Pritchard, 326–334. New York: Routledge, 2011. Surveys a range of issues in the epistemology of memory.

Bernecker, Sven. *The Metaphysics of Memory*. New York: Springer, 2008. Provides both an up-to-date defense of the causal theory of memory and a detailed discussion of alternative theories.

Bloch, David. "Ancient and Medieval Theories." In *Sourcebook for the History of the Philosophy of Mind*, edited by Simo Knuuttila and Juha Sihvola, 205–221. Dordrecht, Netherlands: Springer, 2014. An overview of ancient and medieval theories of memory.

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

Burnham, W. H. "Memory, Historically and Experimentally Considered. I. An Historical Sketch of the Older Conceptions of Memory." *American Journal of Psychology* 2, no. 1 (1888): 39–90. An overview of historical conceptions of memory.

Cheng, Sen, and Markus Werning. "What Is Episodic Memory If It Is a Natural Kind?" *Synthese* 193, no. 5 (2016): 1345–1385. Discusses the nature of episodes in episodic memory.

Clark, Andy, and David Chalmers. "The Extended Mind." *Analysis* 58, no. 1 (1998): 7–19. The classic defense of the extended mind hypothesis, including much-discussed cases of external memory.

Conway, Martin A. "Memory and the Self." *Journal of Memory and Language* 53, no. 4 (2005): 594–628. Presents the "self memory system" framework for autobiographical memory.

De Brigard, Felipe. "Is Memory for Remembering? Recollection as a Form of Episodic Hypothetical Thinking." *Synthese* 191, no. 2 (2014): 155–185. Provides a version of the simulation theory of memory.

De Brigard, Felipe. "The Nature of Memory Traces." *Philosophy Compass* 9, no. 6 (2014): 402–414. An accessible discussion of alternative conceptions of memory traces.

Draaisma, Douwe. *Metaphors of Memory: A History of Ideas about the Mind*. Cambridge: Cambridge University Press, 2000. Describes the history of the varied metaphors by means of which we have understood memory.

Frise, Matthew. "The Epistemology of Memory." *Internet Encyclopedia of Philosophy*. 2015. http://www.iep.utm.edu/. Includes a detailed discussion of the epistemic theory of memory.

Hazlett, Allan. "The Myth of Factive Verbs." *Philosophy and Phenomenological Research* 80, no. 3 (2010): 497–522. Argues against factivity.

Hoerl, Christoph. "Remembering Events and Remembering Looks." *Review of Philosophy and Psychology* 5, no. 3 (2014): 351–372. Argues that remembering is distinguished from imagining by the subject's intentions.

Hume, David. *A Treatise of Human Nature*, edited by L. A. Selby-Bigge, revised by P. H. Nidditch, 2nd ed. Oxford: Clarendon Press, 1975. Contains Hume's theory of memory.

Intraub, Helene, Rachel S. Bender, and Jennifer A. Mangels. "Looking at Pictures but Remembering Scenes." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18, no. 1 (1992): 180–191. Discusses research on boundary extension.

Kania, Andrew, ed. *Memento*. Philosophers on Film. London: Routledge, 2009. A collection of philosophical essays on the film *Memento*.

Klein, Stanley B. "What Memory Is." *Wiley Interdisciplinary Reviews: Cognitive Science* 6, no. 1 (2015): 1–38. An in-depth, philosophically informed discussion of episodic memory in relation to other forms of memory, including the role of the feeling of pastness.

Lackey, Jennifer. "Memory as a Generative Epistemic Source." *Philosophy and Phenomenological Research* 70, no. 3 (2005): 636–658. Discusses the role of defeaters in memory.

Loftus, Elizabeth R. *Eyewitness Testimony*. 2nd ed. Cambridge, MA: Harvard University Press, 1996. An accessible discussion of reconstruction in eyewitness memory.

Martin, C. B., and Max Deutscher. "Remembering." *Philosophical Review* 75, no. 2 (1966): 161–196. The classic statement of the causal theory.

McDonough, Jeffrey K. "Hume's Account of Memory." *British Journal for the History of Philosophy* 10, no. 1 (2002): 71–87. Discusses different interpretations of Hume's criteria for memory.

Michaelian, Kourken. "Generative Memory." *Philosophical Psychology* 24, no. 3 (2011): 323–342. Argues for the compatibility of construction with genuine remembering.

Michaelian, Kourken. "Is Memory a Natural Kind?" *Memory Studies* 4, no. 2 (2011): 170–189. Argues for the disunity of different forms of memory.

Michaelian, Kourken. *Mental Time Travel: Episodic Memory and Our Knowledge of the Personal Past*. Cambridge, MA: MIT Press, 2016. A detailed defense of the simulation theory of remembering, including discussion of the phenomenological character of episodic memory.

Michaelian, Kourken, and John Sutton. "Distributed Cognition and Memory Research: History and Current Directions." *Review of Philosophy and Psychology* 4, no. 1 (2013): 1–24. Reviews the debate over external memory and the extended mind.

Moon, Andrew. "Remembering Entails Knowing." *Synthese* 190, no. 14 (2013): 2717–2729. Defends the epistemic theory of memory.

Pears, David. *Hume's System: An Examination of the First Book of His Treatise*. Oxford: Oxford University Press, 1990. Situates Hume's theory of memory in the context of his larger philosophy and includes a helpful discussion of the vivacity criterion for memory.

Rice, Heather J. "Seeing Where We're At." In *The Act of Remembering*, edited by John H. Mace, 228–258. Malden, MA: Wiley-Blackwell, 2010. Reviews research on field and observer perspective in memory.

Rupert, Robert D. "Memory, Natural Kinds, and Cognitive Extension; or, Martians Don't Remember, and Cognitive Science Is Not about Cognition." *Review of Philosophy and Psychology* 4, no. 1 (2013): 25–47. Attacks the extended mind hypothesis, focusing on the case of external memory.

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210

Shanton, Karen, and Alvin Goldman. "Simulation Theory." *Wiley Interdisciplinary Reviews: Cognitive Science* 1, no. 4 (2010): 527–538. Sketches a version of the simulation theory of memory.

Squire, Larry R. "Memory and Brain Systems: 1969–2009." *Journal of Neuroscience* 29, no. 41 (2009): 12711–12716. Reviews our current knowledge of the memory systems of the brain.

Suddendorf, Thomas, and Michael C. Corballis. "The Evolution of Foresight: What Is Mental Time Travel, and Is It Unique to Humans?" *Behavioral and Brain Sciences* 30, no. 3 (2007): 299–313. A key article on mental time travel; the commentaries (in the same issue) are also very much worth reading.

Sutton, John. "Exograms and Interdisciplinarity: History, the Extended Mind, and the Civilizing Process." In *The Extended Mind*, edited by Richard Menary, 189–225. Cambridge, MA: MIT Press, 2010. Argues for complementarity-based rather than parity-based extended cognition, focusing on memory.

Sutton, John. *Philosophy and Memory Traces: Descartes to Connectionism*. Cambridge: Cambridge University Press, 1998. A classic discussion of memory traces in the history of philosophy.

Sutton, John, and Kourken Michaelian. "Memory." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Winter 2015. http://plato.stanford.edu/archives/win2012/entries/memory/. An advanced reference article discussing memory and pastness, the memory criterion for personal identity, the objects of memory, and many other questions.

Tulving, Endel. "Concepts of Memory." In *Oxford Handbook of Memory*, edited by Endel Tulving and Fergus I. M. Craik, 33–43. Oxford: Oxford University Press, 2000. An accessible discussion of distinctions among different kinds of memory, including nondeclarative memory.

FILM

Memento. Dir. Christopher Nolan. 2000. A man with severe amnesia makes use of external substitutes for memory.

Total Recall. Dir. Len Wiseman. 2012. A secret agent discovers that false memories may correspond to his past more closely than do his apparently real memories. A remake of a 1990 film directed by Paul Verhoeven, itself based on Philip K. Dick's 1966 short story "We Can Remember It for You Wholesale."

COPYRIGHT 2017 Macmillan Reference USA, a part of Gale, Cengage Learning WCN 02-200-210