Chapter One

Memory without a Trace

One of the most persistent conceptual errors in philosophy, psychology, and neurophysiology is the attempt to explain memory by means of memory traces (sometimes called "engrams"). The underlying problems are very deep and difficult to dispel, and as a result, trace theories are quite seductive. In fact, in the cognitive sciences, this approach to memory is ubiquitous and is almost never seriously questioned.[1] If doubts are raised at all, they typically concern how trace mechanisms are implemented or what the physical substrate of traces might be, not whether something is profoundly wrongheaded about the very idea of a memory trace. Moreover, positing memory traces is one aspect of a larger explanatory agenda that prevails in the behavioral sciences—namely, the tempting but ultimately fruitless strategy of explaining human behavior as if it's emitted by, and wholly analyzable in terms of, processes occurring within an agent (typically, inside the agent's brain). And one reason that agenda is so difficult to overturn is that in order to present a viable alternative, one must outline a very different approach to the explanation and understanding of human behavior.

Not surprisingly, that last task is a big one, and I'll make at least a modest introductory stab at it in chapter 3. For this chapter, I have an even more modest (though no less important) goal—namely, to summarize the main reasons for thinking that the concept of a memory trace is not simply useless but actually nonsensical. At the end of the chapter, I'll also show, only briefly, how analogous concepts have crept insidiously into various areas of parapsychological theorizing, especially in connection with the evidence for postmortem survival—for example, speculations about cellular memory in transplant cases and genetic memory in reincarnation cases.[2]

WHY TRACES?

Suppose I meet my old friend Jones, whom I haven't seen in twenty-five years. How is it, we wonder, that I'm able to remember him? Many believe that I couldn't possibly remember Jones without there being something *in* me, a trace (presumably a modification in my brain) produced in me by my former association with Jones. Without that trace, that persisting structural modification in my brain, we'd apparently have causation over a temporal gap. We'd have to suppose that I remember Jones now simply because I used to know him. And to many, that looks like magic. How could something twenty-five years ago produce a memory now, unless that twenty-five-year gap is somehow bridged? So when I remember Jones after twenty-five years, we're tempted to think it's because something in me now closes that gap, linking my present memory to my past acquaintance with Jones.

Now parenthetically, I have to say that it's at least controversial (and in many instances rather naive) to suppose there's something wrong with the idea of causation over a temporal gap. Gappy causation is a problem only on the assumption that the only real causes are proximate causes (i.e., that cause and effect must be spatiotemporally contiguous). But that's a thread I can't pursue here. Positing memory traces is problematic enough quite apart from its underlying questionable picture of causation.

So, let's return to the motivation for asserting the existence of memory traces. Notice that traces aren't posited simply to explain how I happen to be in the particular states we identify as instances of remembering—for example, my experiencing a certain mental image of Jones. They're supposed to explain how memory is *possible* in the first place. The idea is that without a persisting structural modification in me caused by something in my past—in this case, presumably, a physiological representation of Jones—no state in me *could* be a memory of Jones. So if after twenty-five years I have a mental image of Jones, the only way that image could count as a memory of Jones would be if it had the right sort of causal history. And the right sort of causal history, allegedly, is one that spatially and temporally links my present experience with my past acquaintance with Jones. So my image of Jones counts as a memory of Jones only if (1) there's a trace in me, caused by my previous acquaintance with Jones, and (2) the activation of that trace is involved in producing my present image of Jones. So mental images of Jones might be possible without that sort of causal history, but they wouldn't then be instances of remembering.

History has proved that this general picture of remembering is initially very attractive. But it gets very ugly very quickly, as soon as one asks the right sorts of questions. (In my view, this is where philosophy is most useful and often the most fun: showing how claims which seem superficially plausible crumble as soon as their implications or presuppositions are exposed.)

What eventually becomes clear is that the idea of memory as involving *storage* is deeply mistaken and that the mechanism of storage, memory traces conceived as representations of some kind, can't possibly do the job for which they're intended. This is actually an enormous topic and one of the most interesting subjects in the philosophy of mind. But since this issue is both vast and only part of what I want to discuss, I can't do more here than outline a few of the problems with the concept of a memory trace and indicate where one might look for additional details.[3]

MORE PRELIMINARIES

The first thing to note is that the problems with the concept of a memory trace are *hardware independent*. It doesn't matter whether traces are conceived as mental or physical, or more specifically as static, dynamic, neurological, biochemical, atomic, subatomic, holographic (à la Pribram), nonspatial mental images, or (as Plato suggested) impressions in wax. No matter *what* memory traces are allegedly composed of or how they're purportedly configured, they turn out to be impossible objects. Memory-trace theory requires them to perform functions that nothing can fulfill. So my objections to trace theory have nothing to do specifically with the fact that those theories are typically physiological or physical. Rather, it's because they're *mechanistic* and (in particular) because the mechanisms they posit can't possibly do what's required of them.

Before getting into details, I must deflect a certain standard reaction among scientists to the sort of criticisms I'm making here. Many have complained to me that as scientists, they're merely doing empirical research, and so it's simply beside the point to argue, a priori, that their theories are unintelligible or otherwise conceptually flawed. However, I'm afraid that this response betrays a crucial naiveté about scientific inquiry. There's no such thing as a purely empirical investigation. Every branch of science rests on numerous, often unrecognized, abstract (i.e., philosophical) presuppositions, both metaphysical and methodological. These concern, for example, the nature of observation, properties, or causation, the interpretation, viability, and scope of certain rules of inference, and the appropriate procedures for investigating a given domain of phenomena. But that means that the integrity of the discipline as a whole hinges on the acceptability of its root philosophical assumptions. If those assumptions are indefensible or incoherent, that particular scientific field has nothing to stand on, no matter how attractive it might be on the surface. And I would say that several areas of science, as a result, turn out simply to be bad philosophy dressed up in obscurantist technical jargon so that the elementary nature of their mistakes remains well hidden. Memory-trace theory is just one example of this. And I'd argue that today's

trace theories of memory, for all their surface sophistication, are at bottom as wrongheaded and simplistic as Plato's proposal that memories are analogous to impressions in wax. In short, I'd say they're disguised nonsense.

Two more disclaimers, before outlining my objections to trace theory. First, when I say that the concept of a memory trace is nonsensical or that trace theory is conceptually naive in certain respects, I'm not saying that trace theories—or the scientists who hold them—are stupid. To say that a proposal or concept is nonsensical or incoherent is simply to say it makes no sense. Now although the world isn't suffering a shortage of stupidity, not all nonsense is stupid. In fact, the most interesting nonsense is *deep* nonsense, and it's something which can all too easily deceive even very smart people. That's because the problematic assumptions are buried well below the surface and require major excavation.

Second, I've learned over the years that when I outline my objections to trace theory, many hear me as suggesting that the brain has nothing to do with memory. I'll say a bit more about this later, but for now I'll just note that I'm saying nothing of the kind. Although evidence for postmortem survival *would* seriously challenge this, we can overlook for now complications to all physiological cognitive theories posed by the evidence for postmortem survival and restrict our attention to embodied humans. In those cases, clearly, the capacity to remember is causally dependent not simply on having a functioning brain, but probably also on changes to specific areas of the brain. However, it's one thing to say that the brain *mediates* the capacity to remember and another to say it *stores* memories. The former view (more likely the correct one) takes the brain to be an instrument involved in the expression of memory; the latter view turns out to be deeply unintelligible.

THE HORNS OF A DILEMMA

So why is the concept of a memory trace fundamentally nonsensical? Let's begin with an analogy drawn from John Heil's outstanding critique of trace theory.[4] Suppose I invite many guests to a party, and suppose I want to remember all the people who attended. Accordingly, I ask each guest to leave behind something (a trace) by which I can remember them. Let's suppose each guest leaves behind a tennis ball. Clearly, I can't use the balls to accomplish the task of remembering my party guests. For my strategy to work, the guests must deposit something reliably and specifically linked to them, and the balls obviously aren't differentiated and unambiguous enough to establish a link only with the person who left it.

So perhaps it would help if each guest signed his or her own tennis ball or perhaps left a photo of himself or herself stuck to the ball. Unfortunately, this threatens an endless regress of strategies for remembering who attended my

party. Nothing reliably (much less uniquely and unambiguously) links the signature or photo to the guest who attended. A guest could mischievously have signed someone else's name or left behind a photo of another person. Or maybe the signature was illegible (most are), or perhaps the only photo available was of the person twenty-five years earlier (e.g., when he still had hair, or when he had a beard, wore eyeglasses, and was photographed outdoors, out of focus, and in a thick fog), or when he was dressed in a Halloween costume or some other disguise.

But now it looks like I need to remember in order to remember. A tennis ball isn't specific enough to establish the required link to the person who left it. What the situation requires is an *unambiguous representational calling card*, and the tennis ball clearly doesn't do the job. So we supposed that something else might make the tennis ball a more specific link—a signature or a photo. That is, we tried to employ a secondary memory mechanism (trace) so that I could remember what the original trace (the tennis ball) was a trace of. But the signature and photo are equally inadequate. They, too, can't be linked unambiguously to a specific individual. Of course, if I could simply *remember* who wrote the signature or left behind the photo, then it's not clear why I even needed the original tennis balls. If no memory mechanism is needed to make the connection from photo to photo donor or from illegible signature to its author, then we've conceded that remembering can occur without corresponding traces, and then no trace was needed in the first place to explain how I remember who attended my party. So in order to avoid that fatal concession, it looks like yet another memory mechanism will be required for me to remember who left behind (say) the illegible or phony signature or the fuzzy photo. And off we go on a regress of memory processes. It seems that no matter what my party guests leave behind, nothing can be linked only to the guest who left it. We'll always need something else, some other mechanism, for making the connection between the thing left behind and the individual who left it.

In fact, it seems that the only way to stop the regress is for a guest to leave behind something that is *intrinsically* and exclusively linked to only one individual. That's why Wolfgang Köhler, for example, proposed that traces must be *isomorphic* with (i.e., inherently and structurally similar to) the things of which they're traces—that is, the things they represent.[5] But what Köhler and others have failed to grasp is that this kind of intrinsic connection is impossible, because nothing can function in one and only one way. As I'll argue shortly, this is especially clear when the function in question is one of representation or meaning. Nothing can represent unambiguously (or represent one and only one thing); representing is not something objects can do all by themselves; and representation can't be an intrinsic or inherent relation between the thing represented and the thing that represents it.

Interestingly, although Köhler failed to see why trace theory is doomed to fail, he was remarkably clear about what trace theory requires. Köhler understood that a major hurdle for trace theory is to explain trace *activation*—that is, how something present triggers my trace of Jones rather than the trace of someone else. And that's a serious problem, because what triggers a memory (or activates a trace) can be quite different from what established it in the first place. So Köhler wrote,

> recognition . . . means that a present fact, usually a perceptual one, makes contact with a corresponding one in memory, a trace, a contact which gives the present perception the character of being known or familiar. But memory contains a tremendous number of traces, all of them representations of previous experiences which must have been established by the processes accompanying such earlier experiences. Now, why does the present perceptual experience make contact with the *right* earlier experience? This is an astonishing achievement. Nobody seems to doubt that the *selection* is brought about by the similarity of the present experience and the experience of the corresponding earlier fact. But since this earlier experience is not present at the time, *we have to assume that the trace of the earlier experience resembles the present experience, and that it is the similarity of our present experience (or the corresponding cortical process) and that trace which makes the selection possible*.[6]

By the way, this passage reveals another serious limitation of trace theory, one I can only mention in passing here. If trace theory has any plausibility at all, it seems appropriate only for those situations where remembering concerns past *experiences*, something which apparently could be represented and which also could resemble certain triggering objects or events later on. But we remember many things that aren't experiences at all, and some things that aren't even past—for example, the day and month of my birth, the time of a forthcoming appointment, that the whale is a mammal, the sum of a triangle's interior angles, the meaning of "anomalous monism." Apparently, then, Köhler's point about trace activation and the need for similarity between trace, earlier event, and triggering event, won't apply to these cases at all. So even if trace theory were intelligible, it wouldn't be a theory about memory generally.

In any case, trace theory is not intelligible, and Köhler's observation reveals why. To avoid the circularity (and regress) of positing the ability to remember in order to explain my ability to remember (e.g., by requiring further trace mechanisms to enable the previous trace do its job), we must suppose that some trace uniquely and unambiguously represents or connects to the original experience. And because unambiguous representation is an impossible process, trace theory is caught between two fatal options. I'll explain in a moment why unambiguous representation is impossible, but

first, we need to observe that the tennis ball/party example hides a further complication noted in the passage from Köhler.

Traces are usually supposed to be brain processes of some sort, some physiological representation produced, in this case, by a party guest. But what *activates* this trace later can be any number of things, none of which need to resemble the experience, object, or event that produced the original trace. Suppose Jones attended my party. Trace theory requires my experience of Jones at the party to produce a representation in me of Jones (or my experience of him) so that I can later remember that he was at the party. But what will eventually activate that trace? It could be Jones himself, or an image of Jones, or the lingering smell of someone's cologne, or a telltale stain on the carpet, or perhaps someone asking, "Who was at the party?"

Of course, some of these potential triggering objects or events might plausibly be said to resemble the thing that originally produced the trace. But how can (say) the smell of cologne, a stain, or the words "Who was at the party?" trigger the trace of Jones created by his presence at the party? These things aren't obviously similar to Jones himself. If we posit another memory *mechanism* to explain how I draw the connection between the cologne and Jones (e.g., he may have worn it, spilled it, or simply talked about it) or how the question "Who was at the party?" leads me to the right party and not some other party, *or even how I remember what the word "party" means*, we're starting a regress of memory mechanisms. But if we say it's because I can simply *remember* who wore (or perhaps mentioned) the cologne, who stained the carpet, or who my party guests were, or what "party" means, then we're still reasoning in a circle. We're still explaining memory by appealing to the ability to remember. Moreover, if I can remember these things without some further trace, then we didn't need a trace in the first place to explain my ability to remember that Jones was at the party. However, if we follow Köhler's lead, then we have to assert some kind of intrinsic similarity or resemblance, some kind of psychophysical *structural isomorphism*, between three—potentially quite different—things: the original experience or event, the trace produced on that occasion, and the subsequent triggering events.

Furthermore, if (like me) you believe that the meanings of sentences or words aren't things that have a structure, something whose parts and relations between them can correspond to another structure in the brain (or somewhere else), a trace theory of memory can't appeal to a system of representations and structural isomorphism or similarity to explain how one remembers the meanings of words. But then trace theory has to be *completely mute* on the question: How does the sentence "Who was at the party?" trigger my memory that Jones was at the party?

If nothing else, these considerations should make you suspicious that an inner (brain) representation of Jones at the party can be isomorphic both to Jones (or my experience of him) and to the innumerably many and quite

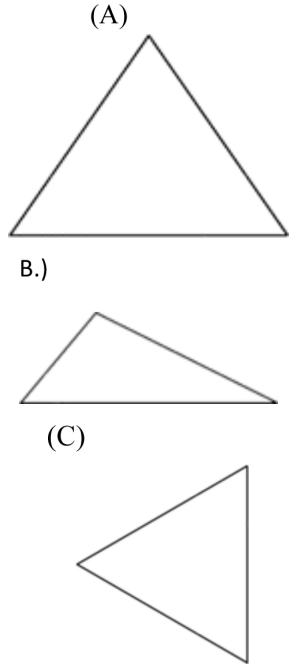
different things that can later activate the trace—for example, a particular scent or a sequence of sounds. What kind of similarity could this be? The answer is that it can't be any kind of similarity and that Köhler's proposal is literally meaningless. As tempting as it is to continue for a while enumerating the problems with trace theories, I'll restrict myself now to two more points, to explain perhaps the deepest confusion underlying these theories.

The first problem is with the very idea of structural isomorphism. The term "structural isomorphism" sounds impressive and scholarly, but in trace theories, the appeal to structural isomorphism is really just the appeal to an *inherent similarity* between two things, *determined solely by their respective structures*. It's merely a kind of copying, and perhaps if trace theorists spoke only of copying rather than isomorphism, their theories would appear as silly as they really are. That's why nobody takes seriously the theoretically identical position that Plato proposed—namely, that memories are like impressions in wax. It sounds much more impressive to speak instead of isomorphism, and it's also much more effective than speaking of wax impressions, because it drives the confusions and theoretical silliness underground. But the unavoidable nonsense of trace theory remains, and the crucial point is this: traces must be produced in a way that relates them structurally to the things of which they're traces, and they must be activated only by things having the right underlying structure. Moreover, that activation must be determined solely by intrinsic relations between the structures of the trace and the things that activate them. Otherwise, we'd need another mechanism to explain how the *right* trace is activated in the presence of a trigger that could just as well have been isomorphic with (or mapped onto) something else. And that raises the circularity or regress problem noted earlier.

But the alternative, inherent similarity, makes no more sense than saying that a square is a circle. Inherent similarity is a *static* relation obtaining only between the similar things. And it must hold between those things *no matter what*. If, for example, context could alter whether two things count as similar, then those things are not similar merely in virtue of intrinsic relations holding between their respective structures. But that's why intrinsic similarity is nonsense. Context *does* matter; in fact, it's indispensable. Things are never similar solely in virtue of static relations between them or in virtue of properties inherent in them. Things must *count* or *be taken as* similar or dissimilar relative to some context of inquiry and criteria of relevance. For example, the movements of an elephant aren't inherently similar or dissimilar to those of a flea. They might count as similar in a situation where the size of an organism isn't relevant but dissimilar in a context where size is a major concern. Or suppose I try to tell the same joke I heard someone tell the day before. Is the joke I told similar or not to the one I heard earlier? Obviously, it depends on what's relevant to our answering that question, and no criteria of relevance are inherently privileged over the others. Depending on the

situation, we might focus on whether my joke made the audience laugh, whether the words were exactly the same, whether they were delivered at the same speed, or with the same accent or timing, or with the same inflection, or whether my voice had the same timbre as that of my predecessor. Although cognitive scientists and memory theorists seem to ignore it all the time, the point is painfully clear: similarity exists *only* with respect to variable and shifting criteria of relevance. It can only be a dynamic relation holding between things at a time and within a context of needs and interests.

Another example, this time from geometry, should make the point even more clearly. Consider the triangle (A) in figure 1. Then, compare it to the geometric figures (B)–(E). Now consider the question: To which of the last four figures is (A) similar? The proper response to that question should be puzzlement; you shouldn't know how to answer it. Without further background information, without knowing what matters in our comparison of the figures, the question has no answer at all. Mathematicians recognize this, although instead of the term "similarity" they use the expression "congruence." In any case, mathematicians know that in the absence of some specified or agreed-upon rule of projection or function for mapping geometric figures onto other things, no figure is congruent with (similar to) anything else.



D.)

E.)

Mathematicians recognize that there are different standards of congruence appropriate for different situations. But no situation is *intrinsically basic*, and so no standard of congruence is inherently privileged or more fundamental than others. For example, engineers might sometimes want to adopt a fairly strict mapping function according to which (A) is congruent only with other figures having the same interior angles and the same horizontal orientation. In that case, (A) would be congruent with none of the other four figures. Of course, only in very specialized contexts are we likely to compare figures with respect to their horizontal orientation. In many situations it would be appropriate to adopt a different standard of congruence, according to which sameness of interior angles is all that matters. And in that case we'd say that figures (A) and (C) are congruent but that (A) is not congruent with the other figures. However, there's also nothing privileged about sameness of interior angles. Perhaps what matters is simply that (A) is congruent with any other three-sided enclosed figure, in which case we could say it's congruent with the three triangles (B)–(D) but not with the rectangle (E). But even that criterion of congruence can be modified or supplanted. Mathematicians have rules of projection that map triangles onto any other geometric object, but not (say) to apples or oranges. Of course, the moral here is obvious. If simple geometric figures aren't intrinsically similar—that is, if they count as similar only against a background of assumptions about which of their features matter (i.e., are relevant), then we certainly won't find intrinsic similarity with much more complex objects—in particular, memory traces and the various objects or events that allegedly produce and activate them.

But maybe you're still not convinced. Perhaps you think that there *is* a fundamental principle of congruence for this geometric example. You might think that, first and foremost, (A) is similar to just those figures with sides of *exactly* the same length, the same horizontal orientation, and exactly the

same interior angles. And perhaps you'd want to call that something like "strict congruence (or identity)." But there are at least three serious problems with that position.

First, even if this sort of congruence counted as more fundamental than other forms of geometric similarity, that could only be in virtue of a kind of historical accident. The primacy of that standard of congruence would reveal more about us, our conventions and values—in short, what merely happens to be important to us—than it does about the figures themselves. In fact, it's a standard appropriate for only a very narrow range of contexts in which we consider whether things are similar. Second (and as an illustration of that first point), it's easy to imagine contexts in which two triangles have exactly the same interior angles, horizontal orientation, and sides but don't count as similar. If we're interior designers, for example, it might also matter whether the triangles are of the same color, or whether they're placed against the same colored background, or whether they're made of the same material. If we're graphic artists, it might matter whether the triangles were both original artworks or whether one was a print. Or if we're librarians or archivists, it might matter whether the triangles occur on the same page of different copies of the same book. And third, even if we could decide on some very strict sense of congruence (or identity) that would count as privileged over all other forms of similarity, it would be useless in the present context. Memory traces are never strictly identical either with the things that produce them or with the things that activate them. The looser and more complex forms of similarity at issue in trace theories are classic examples of the sorts of similarities that can't possibly be inherent, static relations between things.

And as if that weren't enough, another aspect of this general confusion about similarity is the requirement that traces and other things have intrinsic or inherent *structures*—that is, some context-independent parsing into basic elements. Because isomorphism (mapping) is tied to structural elements of the isomorphic things, that's a necessary condition for intrinsic isomorphism to hold between the trace and the things it represents. After all, if what counted as structure depended on context—that is, if a trace could just as well have been parsed differently and assigned alternative structures—then it could be mapped onto (or count as similar to) different things. And, unfortunately for trace theory, objects and events can always be parsed in an indefinite number of ways, and whatever parsing we select can only be conditionally, and never categorically or intrinsically, appropriate. We always determine a thing's components relative to a background against which certain features of the things (but not others) count as relevant. But then it's only against shifting and nonprivileged background criteria of relevance that we take two things to have the same structure; they're never isomorphic *simpliciter*—that is, intrinsically or inherently.

So the trace theorist's inevitable appeal to privileged, inherent structures and intrinsic mappings is literally absurd. It's on a par with claiming that a pie has a basic or privileged division into slices or elements—that is, a context-independent answer to the question "How many pieces are there to this (unsliced) pie?" as if the number of potential pie eaters were irrelevant to our answer. Similarly, it's as absurd as claiming that there's an absolutely context-independently correct and privileged answer to the questions "How many events were there in World War II?" and "How many things are in this room?" If we consider the first of those questions, it's clear that our answer depends on how much of a bird's-eye view we're taking on World War II, and that depends entirely on the purpose of the discussion in which the question arises. In a broad discussion of military history generally, it might be enough to parse the war into just the European and Pacific campaigns. But in more specialized discussions, finer-grained parsings are likely to be more appropriate—say, into particular battles alone, or battles plus relevant meetings and decisions of world leaders and military commanders. Independent of some such context, the question "How many events were there in World War II?" has no answer at all; there's simply no basis for dividing it into certain parts rather than others.

CONFUSIONS ABOUT REPRESENTATION

The appeal to inherent similarity or structure is merely a specific form of a more pervasive problem in the so-called cognitive sciences—namely, confusions about and equivocations on the term "representation." Traces are supposed to represent their causes, the objects, events, or experiences that produced them, and they must be internally and structurally differentiated in ways that correspond to the different things we remember. In other words, the trace of (or the internal state that represents) Jones at the party must differ structurally from the trace of (or the internal state that represents) Smith at the party, or Jones at some other party, or my dissertation defense, or the joke someone told the previous evening. And those traces (or internal states) must differ structurally from one another in ways corresponding to the respects in which their causes differ structurally. After all, if my trace of last night's joke wasn't uniquely and structurally distinct from my trace of Jones at the party, then trace theorists would have no way to explain how activation of a trace produces one memory rather than another—and the right one at that. So trace theory is one version of the general view that particular mental states are caused by (or are identical to) certain corresponding distinct internal brain states, and that what those different internal states *are* (i.e., what they represent) is explainable wholly in terms of their distinctive structural features. At this point, cognitive scientists typically do a lot of hand waving and say

something like, "We may not currently know all the details, but presumably some psychologist in the future (or perhaps God) would be able to look inside our heads and know, from the way we're configured, what we're thinking."

However, this general picture rests on the utterly false assumption that a thing's representational properties can be determined solely by its structural or topological features. I've examined this error in considerable detail elsewhere.[7] For now, a few brief remarks will have to suffice.

To see what's wrong, we need to appreciate that *anything can represent anything*. In fact, a thing's representational options are limited only by the situations into which it can be inserted. And that set of situations is as indefinite and vast as the set of possible twists and turns human life can take. But if that's the case, then what something represents can't simply be a function of how it's configured. Things must be *made* to represent or mean something. Suppose I'm trying to teach a child the alphabet. I show him a picture of a dog and I say, "*D* is for dog." In that case, we might say that the picture represents the class of dogs. But I could have said, "*C* is for collie," and in that case, the picture would have represented a subset of the set of dogs. Similarly, I could have said "*L* is for Lassie," in which case the picture would have represented an even smaller subset of dogs. I could also have said "*Z* is for Ziggie," referring to the child's pet collie. And notice, these changes in what the picture represents have nothing whatever to do with corresponding changes in the arrangement of pixels, or atoms, or anything else in the picture. Those structural features of the pictures remained the same in all cases. What the picture represents depended instead on how it was used.

In fact, the picture's representational properties could be changed even more dramatically. My disgruntled students could make the picture represent me and symbolically express their hostility toward me by using it as a target for darts. Or I could jokingly point to the picture and say, "This was Joan Rivers before plastic surgery." Or suppose I'm trying to give directions to someone without the aid of a map. I could place the picture on a table and say, "This is the shopping center, this [a tuna sandwich] is the hospital, this [my fork] is the access road, and this [a salt shaker] is the water tower."

Of course, contexts in which (say) a sandwich represents a building or in which a picture of a dog represents a distinguished philosopher (or over-the-hill comedienne) are atypical in some respects. But those situations are unusual *only* with respect to what the objects represent. They aren't at all unusual with respect to how representational properties are acquired. And it doesn't matter whether we're talking about images, words, or (say) synaptic connections. In every case (familiar and offbeat), what a thing represents depends ultimately on the way we place it in a situation, against an enormously rich background of needs and interests and both local and global

traditions and assumptions about the way the world is and which things matter. *There are no purely structural or context-independent forms of representation or meaning.* So when it comes to examples like the picture of a dog or the tuna sandwich, the mistake many make is to think that some representational properties—the familiar and apparently default ones—are inherently fundamental and that others are anomalous. That is, they believe that representation in familiar cases is somehow built into or hardwired into the representing objects and that this inherent function simply gets *overridden* in the more unusual cases. But in fact, the familiarity of certain contexts reveals more about us, about our patterns of life and our interests, than it does about the objects themselves. If our form of life were radically different, the ostensibly default or familiar representational properties of objects could change accordingly.

But then, if a brain structure (say) is to represent something past and function as a memory trace, it can't do so solely in virtue of its structural features. Nothing represents or means what it does on topological grounds alone. However, the whole point of Köhler's principle of psychophysical isomorphism (or related hypotheses in the cognitive sciences) is to tie what a thing represents solely to its structure. That was the only way to avoid the equally fatal error of requiring a regress of mechanisms to explain how the original mechanism or state can do its job. So this, too, turns out to be a dead end.

TOKENS AND TYPES

But let's return more explicitly to trace theory. A related and equally unheralded problem with trace theory is its ontology. It posits an entity that's of a radically different kind from the concrete things in the world that are supposed to cause and activate memory traces. And that's a problem because the sort of thing this entity must be is something that many believe is a philosophical fiction. In fact, positing the existence of memory traces is more a philosophical move than a scientific move. Hopefully, one distinction and one more example will make this clear.

Trace theorists have always been tempted to regard traces as kinds of *recordings* of the things that produced them. In fact, some previous influential writings on memory compared traces to tape recordings or grooves and bumps in a phonograph record. The justification for that idea, as we've seen, is that traces must somehow capture essential structural features of the things that both produce and activate them. That's one of the keys to how trace theory is supposed to work. Allegedly, what links together and unifies traces both with their causes and their activators is a common underlying *structure*.

So the issue we must now address is: What sort of thing is this structure? I'll argue that the required structure is an impossible object.

Consider: one of the things I remember is Beethoven's Fifth Symphony (hereafter abbreviated as B5). Modern versions of trace theory require that my memory be explained in terms of a representation of B5, stored in some concrete physical form in my brain and (let's say) produced in me by an experience of hearing B5 in the past. This trace must have certain structural or topological properties that link it both to the thing(s) that caused it and also to those which later activate or trigger it. These must also be properties that distinguish the trace of B5 from traces of other pieces of music. So presumably, this B5 trace was produced by and captures specific features of a performance I heard of B5 which will enable activation of the B5 trace in the presence of subsequent items sharing those specific features.

But which features might these be? Tempo, rhythm, pitch, length of notes, instrumental timbre, dynamic shadings? You'd think so if my trace of B5 was produced by and represents or records a B5 performance and also if that trace is to differ (say) from my trace of Beethoven's Fourth (B4) or even "Yankee Doodle." But I (like many others) can remember B5 by recognizing a wide variety of musical performances as *instances* (or as philosophers would put it, *tokens*) of B5. And these tokens can differ from one another and from the original trace-producing instance of B5, with respect to *any* features of that original event. Even wild parodies of B5 are instances or tokens of B5. That's why I can tell what they're parodies *of*—that is, that they're B5 parodies. Obviously, that feat of identification is one form of remembering B5. For instance, I could recognize B5 when certain notes are held for an unusually long time, or when it's played with elaborate embellishments, or with poor pitch and many mistakes by an amateur orchestra. In fact, I could recognize truly outlandish musical events as instances of B5—for example, when it's played extremely slowly or rapidly, or with tempo changing every bar, or with arbitrary notes raised a major sixth, or when it's played with inverted dynamics or played only on kazoos, banjos, or tubas. Similarly, I could recognize a series of percussive taps as a pitch-invariant version of the opening bars of B5.

So what is it that the B5 trace has in common with the concrete events that can cause and trigger it? As we'll see, it must be a very unusual sort of entity. Whereas the remembered and triggering events are (typically) concrete instances of B5 (e.g., performances of one kind or another), and whereas the trace itself is also a specific, concrete thing—that is, some kind of persisting modification of the person (e.g., a brain state), this common unifying element must be a relentlessly abstract object—what philosophers call a *type*. Moreover, since memory-triggering instances of B5 can differ from the original trace-producing event with respect to *any* of the original event's features (e.g., rhythm, dynamics, timbre, tempo, pitch, absence or presence

of embellishments, etc.), and since the B5 trace is presumably an even more radically different kind of *version* of B5 (say, a neurological version that itself has no pitch, tempo, timbre, etc.), the structure that they allegedly share must be so abstract that it contains *none of the concrete musical features* found in the events that produced it (e.g., precise rhythm, pitch, etc.). In fact, it can't have specific features found in any *possible* version or embodiment of B5.

Remember that the B5 trace is supposed to provide a unifying permanent structural link between the past's original trace-producing B5 token and all possible subsequent B5 tokens that activate the B5 trace. Trace theory requires that all those events and the trace are B5s because they share a common underlying structure. That's how Köhler's principle of psychophysical isomorphism works. It's supposed to help explain how the right trace, the B5 trace, gets picked out and activated by an event that could, in principle, be mapped onto (linked structurally to) something else. Although we've seen that this explanatory strategy relies on the incoherent notion of intrinsic similarity (or else falls victim to an endless regress of memory mechanisms), the outlines of the strategy are clear enough. We're supposed to believe that certain events activate the B5 trace because they share a deep structure with that trace, and also that the trace is a B5 trace because it shares that same structure with the instance of B5 that produced it.

But in that case, it's reasonable to ask the trace theorist: Which *specific* features, exactly, might these various tokens of B5 have in common—that is, which features count as parts of the underlying common structure? Notice, that question will be difficult to answer so long as any specific feature of any instance of B5 can be absent from some other token. This is easy to see even if we consider only musical performances. If one B5 token is at a certain pitch, or volume, or tempo, or whatever, some other token might be at a different pitch, volume, tempo, and so on. And then, of course, if the B5 trace is a brain state (as trace theorists usually suppose), it has neurological (biochemical, or whatever) properties that *represent* musical properties. It has no musical properties (such as pitch, tempo, etc.) itself. So it appears that the hypothetical common deep structure linking the B5 trace to all concrete instances of B5 has none of the specific features found in any actual instance of B5. That's why trace theorists have no choice but to posit a common *abstract* type (e.g., a B5 type) linking the indefinitely many and different possible B5 event tokens. The memory trace, the original trace-producing event, and later triggering events would all *exemplify* or embody this type, in virtue of sharing that type as their common underlying structure.

To see this more clearly, let's review some issues discussed earlier. Suppose that the only performance of B5 I ever heard—and thus, the performance that produced my trace—was a conventional and accurate modern orchestral performance. How, then, is that B5 trace picked out and activated

by my hearing something radically different—for example, a thoroughly novel and (thank heaven) once-in-a-lifetime accordion-only performance of B5 played at quarter speed, transposed to another key, with many wrong notes, and embellished as only accordion players can? That is, how do I recognize that this nightmare musical event is a B5 and not (say) an anthem played at the unveiling of a central European monument? How does this nightmare musical event activate the B5 trace rather than one or more of the myriad other events we could consider similar to it? If we try to prevent a regress by saying that I can simply recognize that the accordion-only performance is an instance of B5, then we don't need to posit the trace of B5 at all. We've conceded that I can remember and thereby recognize B5 without recourse to a B5 trace. Again, that's one reason Köhler and others appealed to structural isomorphism—intrinsic similarity—between the original event that produced the trace, the memory trace itself, and the triggering event that activates the trace. As we've already seen, that's a fatal move because the very idea of intrinsic similarity is conceptually confused. Moreover, it clearly wouldn't work anyway in the sorts of cases we're considering now, where original and triggering events differ dramatically in their concrete properties. That's supposed to be one of the attractions of appealing instead to properties of a common abstract type (e.g., a B5 type) that cuts right through these concrete variations and links the different event tokens to one another and to the trace.

But now look at what's happened. We've seen that the common element linking all B5 tokens as B5s isn't something that has any particular features of any particular performance or instance of B5. But it's still supposed to be a kind of thing, *the structure* that all B5 tokens share and in virtue of which they're B5s rather than B4s or other tokens. However, that structure has to have *some* features in virtue of which it's a B5 structure and not that of (say) Beethoven's Fourth, the "Waldstein" Sonata, or "Yankee Doodle." But it can't have features found in any specific instances or tokens of B5, because for whatever feature we specify, some other instance of B5 might lack it. But then no specific feature of an instance of B5 can be necessary for something's being a B5. So although it can't be anything like any actual B5 performance, with specific pitches, dynamics, rhythm, and so on, the common B5 structure—that is, *the* thing that all B5s have in common—somehow needs to have features necessary for its being *the* B5 structure but also distinctive to its being the B5 (rather than, say, a B4) structure. In fact, it must *inherently* be the kind of structure it is. And it must inherently be a B5 structure, despite lacking all properties that B5 tokens might have.

We can't even say that, whatever sort of properties the B5 structure has, they must at least be—presumably abstract—*musical* features (pitch, rhythm, etc). For one thing (as we've noted), if a B5 trace in the brain counts as a B5, its neurological or biochemical (or whatever) properties would be of an en-

tirely different sort. Morever, it's not clear what an abstract musical feature would be. A determinable property, like *some pitch or other*? Some tempo or other? That wouldn't enable the B5 structure to differ from a B4 structure. Those different common structures must differ with respect to determinate (i.e., actual, specific) properties, not determinable properties.

In short, the common B5 structure must have features *necessary and sufficient* for its being a B5 and not (say) a B4 but without having any specific features regarding pitch, tempo, dynamics, and so on, any of which might be absent from any concrete instance of B5. (Perhaps you can now see why many consider abstract types to be impossible objects.) So what allegedly links the B5 trace to both the remembered and triggering events is a shared, abstract, B5 structure, which—incredibly—is inherently that of a B5 (and not a B4) but which has no specific features of an actual instance of B5, no actual specifications or instances of pitch, dynamics, rhythm, and so on.

To avoid even further embarrassment to the trace theorist, we can conveniently ignore for now cases where what causes me to remember B5 (i.e., what activates the trace) is not another musical event but something that can't even remotely be considered to have the same underlying abstract structure as a concrete instance of B5—for example, a portrait of Leonard Bernstein, a hearing aid, or the question "Can you hum the first few bars of that symphony you heard last week?"

In any case, we've arrived at the point where we see the ultimately nonscientific nature of trace theory. It's committed to the view that a memory trace of B5 and all concrete instances of B5 have a structure that is essential to all things that are instances of B5 but none of the specific features that real, concrete versions of B5, including the trace itself and nightmare versions, can lack. This position is commonly called Platonic essentialism—the view that things are of the same kind in virtue of sharing a common underlying but abstract structure. And that's not a scientific view at all. It's a philosophical view, and a bad one at that.

RECENT MEMORY RESEARCH

A predictable rejoinder to the foregoing arguments against trace theory would be that memory research has progressed considerably since the days of Köhler. Some might even suggest that talk of memory traces is now passé. So perhaps the position I've defended is simply out of date and my arguments just don't apply to current memory theory.

Granted, there's been undeniable progress in thinking about the domain of memory—for example, taxonomic advances in identifying the varieties or types of memory. Likewise, increasingly advanced technology has enabled researchers to probe our neurophysiological systems in unprecedented detail.

Nevertheless, in a crucial respect, recent memory research shows no progress at all and remains defiantly superficial. It takes for granted that some form of storage and retrieval takes place in the brain, whether the putative physiological mechanism is a unitary engram or something distributed or diffuse—say, across a cell assembly, and whether what's stored in the brain is static or dynamic. But it never addresses the fundamental issues of how any physical modification can represent or stand for what is remembered and indeed how it can represent or stand for one thing rather than another. And trace talk is alive and well.

A full review here of the broad spectrum of recent memory work is out of the question, but the brief survey below should make the point handily. Consider, for example, the sorts of proposals that create flurries of excitement within the community of memory researchers. One recent innovation was to revise the long-held view that memories are initially labile and must be stabilized or consolidated before they become long-term memories. The innovation was to argue that long-term memories aren't as firmly rooted as once believed and that "memories, or parts of them, need to be restabilized after their expression in a manner analogous to the initial stabilization process."[8] In particular, some claimed that once memories are retrieved, they become labile again and need to be reconsolidated.[9] Another hotly debated proposal (first aired by Todd Sacktor) concerns the alleged importance of an enzyme named protein kinase M-zeta (or PKMζ) in sustaining long-term memories in the brain.[10] This is even described by Sacktor as "a candidate, persistent enzymatic molecular mechanism for the longterm memory trace."[11] However, more recent work seems to have successfully challenged this alleged role for PKMζ.[12]

Significantly, nowhere in any of this work will you find critical reflection on whether traces are possible objects or whether the concept of memory storage is tenable. In recent work it's simply taken for granted that information is stored somehow in the brain, as if the matter were settled a long time ago and all that we need to do now is to figure out what the correct hardware description of the process is. So that fundamental assumption is never defended or scrutinized, and the problems with it and with its associated reliance on the notion of representation apparently go unrecognized. But then, despite the technical and technological advances of the latest work on memory, that work is ultimately even less sophisticated conceptually than the old work of Köhler. Köhler at least realized that the very idea of storage and retrieval had an ineliminable philosophical component and that the posited memory mechanism—whatever its precise hardware realization—required an appeal to the principle of isomorphism, the defects of which we examined earlier.

As another example, consider how Karim Nader summarizes the thrust of his paper "Memory Traces Unbound" (and for that matter, much of his recent

research): "The idea that new memories are initially 'labile' and sensitive to disruption before becoming permanently stored in the wiring of the brain has been dogma for > 100 years. Recently, we have revisited the hypothesis that reactivation of a consolidated memory can return it to a labile, sensitive state—in which it can be modified, strengthened, changed or even erased!"[13]

Notice, what's at issue here is only whether (or to what extent) the physiological storage of memories in the brain is permanent—not whether there's any kind of storage at all or whether storage of the kind required is even possible. Evidently, it never occurred to Nader that a more profound dogma is that memories can be "stored in the wiring of the brain."

This is by no means an isolated case. In a comprehensive four-volume set from 2008 intended to display the state of the art in research on memory and learning,[14] we find the following, all too typical, passages. (a) "The notion of a physical memory trace, independent of its use . . . is a central presumption in neuroscience."[15] (b) "A neuroscientist cannot help but assume that the knowledge stored in memory continues to exist during time periods when it is not retrieved."[16] Remarkably enough and significantly, these statements appear in the introductory essay to the entire four-volume set.

Some memory researchers believe they've made advances by supposing that there are different, or different kinds of, memory systems, corresponding to different kinds of memory—not simply (say) memory for past events—what is sometimes called episodic and autobiographical memory.[17] For example, John O'Keefe and Lynn Nadel write, "It appears that there are different types of memory, relating perhaps to different kinds of information, and that these are localized in many, possibly most, neural systems."[18] And later, "There is no such thing as the memory area. Rather, there are memory areas, each responsible for a different form of information storage.[19] Again, the process of memory storage is simply taken for granted, and it's uncritically assumed to be due to some kind of localized physiological change in the brain.

Similarly (from the Byrne four-volume set), Nadel concedes that all "memory systems" rely on the same fundamental presupposition: "It makes . . . sense to think about all neural systems as both processing and storing knowledge, with the differences between systems reflecting the nature of the knowledge being processed and stored, and the timescale of that storage."[20]

It's very easy to multiply examples. Raymond Kesner defends "a tripartite attribute-based theoretical model of memory that is organized into event-based, knowledge-based, and rule-based memory systems."[21] But "Each system is . . . mapped onto multiple neural regions and interconnected neural circuits."[22] And what does that talk of mapping mean? Well, for instance, "The event-based memory system provides for temporary representations of incoming data concerning the present" and the "knowledge-based memory

system provides for more permanent representations of previously stored information in long-term memory and can be thought of as one's general knowledge of the world."[23]

But of course, this is just recourse to the old talk of representation and storage. No deep problems are solved by the recourse to multiple memory systems rather than one so long as the idea of physiological storage is retained. So this multiple-system position remains a house of cards. If the strategy for explaining memory of even just past events fails for the fundamental reasons discussed earlier, it undermines all the attempts to ground additional forms of memory "storage" in their respective neurophysiological systems. But even if some forms of memory—say, memory-how (procedural or skill memory)—escape the critique of this chapter, memory theory still suffers a colossal failure if it can't explain memory of past events in terms of storage and representation.

One final example. A recent study, admittedly involving only mice, explored the "optogenetic reactivation of hippocampal neurons activated during fear conditioning,"[24] and the results were taken to indicate that "activating a sparse but specific ensemble of hippocampal neurons that contribute to a memory engram is sufficient for the recall of that memory."[25] It's worth noting, again, how the authors describe the conceptual background against which their study should be viewed. They write, "An important question in neuroscience is how a distinct memory is formed and stored in the brain. Recent studies indicate that defined populations of neurons correspond to a specific memory trace, suggesting a cellular correlate of a memory engram."[26]

But this is simply old-fashioned trace theorizing centering on another candidate du jour for the locus of the trace. So I must reiterate a point made earlier in this essay. It doesn't matter what the hardware account of memory traces is—for example, whether it's static or dynamic, where in the brain it is, or how it's configured or localized. It doesn't even matter whether the trace is physical. The problem with trace theory is that traces, however they're conceived, are required in order to perform a function that no object can fulfill. And trace theory relies on a concept of similarity—intrinsic similarity—that literally makes no sense.

TRACE THEORY IN PARAPSYCHOLOGY

It's unfortunate enough that memory-trace theory is received dogma in the cognitive sciences. Almost no one seems to doubt that memories are somehow stored and encoded in us. So it's not surprising that this picture of memory has found its way to more overtly speculative or frontier areas of science, including parapsychology. No doubt it's very tempting for para-

psychologists to posit trace-like processes in their own theories, because they can then at least appear to be reasoning along scientifically orthodox lines, even if the subject matter itself falls outside the scientific mainstream.

For example, William Roll proposed a "psi structure" theory of postmortem survival, modeled explicitly after memory-trace theory and according to which memory traces are left not simply in individual brains but in our environment as well.[27] Of course, this escapes none of the classic problems of trace theory, because in Roll's view, what certain structures represent (or are similar to) remains unintelligibly tied to inherent features of those structures. This is especially problematic when Roll suggests that an individual mind or personality is a system of such structures. That's no more plausible than saying that we can tell whether a person is thinking about his grandmother just by examining the state of his brain, or that a picture of a dog represents something specific independent of its use in a context. Roll's view requires brain or mental structures to mean or represent something simply in virtue of how they're configured, never mind their dynamic position within an equally dynamic life situation. Roll also proposes explaining ESP as the response to memory traces left on objects by previous guesses. But that seems no more credible than supposing that my ability to remember my party guests is simply a mechanical function of the tennis balls they left behind or the illegible signatures or photos they left along with the balls.

Trace theory also appears in other guises in connection with the evidence for postmortem survival. One is the suggestion that reincarnation cases can be explained in terms of genetic memory. However, I've found no serious researcher making that suggestion. It seems, instead, to be entertained simply as a real possibility, albeit one that can be rejected on empirical grounds.[28] That is, it's treated as if it's an intelligible position that happens merely to be inadequate to the data. Another application of trace theory to survival is the attempt to explain transplant cases by appealing to cellular memory.[29] No doubt the reason it's tempting here to posit genetic or cellular memory traces is that in reincarnation and transplant cases, complex psychological regularities seem to persist in the absence of the usual presumed bodily correlates. So to those for whom it's unthinkable that individuals can remember without their memories being stored somewhere, it might seem reasonable to propose that memories and personality traits can be encoded in a kind of hardware that has nothing to do with the brain. However, since the problems noted earlier with trace theories are hardware independent, it's an insignificant change merely to relocate the traces in different physical systems. It's still untenable to suppose that representation, meaning, or similarity are determined solely by a thing's topological features.

To me, it's interesting that when the usual suspect—the brain—isn't available as the locus of memory storage, some find it inevitable that memories must simply be located in a different place or perhaps in a modified

form. It demonstrates just how deeply mechanistic assumptions have taken root, and in a way, it shows a profound lack of scientific imagination. The situation here closely parallels what happened in response to Karl Lashley's famous experiments in the 1920s.[30] When Lashley found that no matter how much of a rat's brain he surgically removed, trained rats continued to run their maze, some concluded that the rats' memories weren't specifically localized in their brains. Instead, they suggested that the memories were *diffusely* localized, much as information is diffusely distributed in holograms.[31] But to someone not antecedently committed to traditional mechanistic dogma, Lashley's experiments take on a different sort of significance, perhaps similar to that of the evidence for postmortem survival. They suggest that memories aren't located anywhere or in any form in the brain. And more generally, they suggest that the container metaphor (that memories and mental states in general are *in* the brain or in something else) was wrong from the start and also that memories (and mental states generally) aren't *things* or *objects* with distinct spatiotemporal coordinates. Of course, that's what my arguments in the preceding sections were intended to show.

Another variant of this general error emerges in Rupert Sheldrake's suggestion that morphic fields capture the essential structure of developmental forms and even behavioral kinds. Although Sheldrake thought he was escaping the evils of mechanistic theories with his view, in fact he retained the underlying errors of supposing that similarity is an intrinsic structural relation between things and that things of the same kind are of that kind because they share a common underlying structural essence. The claim that behavioral kinds, such as feeding behavior and courtship, can be captured in strictly structural terms is especially implausible. (For a detailed critique of Sheldrake's theory, see chapter 2.)

SUMMING UP

I realize that I'm pretty much a voice in the wilderness on these issues, and I find myself in the unenviable position of having to argue that many prominent and respected scientists actually don't know what they're talking about. I wish there were some other, less fundamentally upsetting, way to undercut trace theories of memory. But I believe that the problems really are that deep and that the theories really are that essentially confused.

However, as long as I'm being antagonistic, I see no compelling reason to stop where I left off. I might as well finish with a brief obnoxious coda. As I see it, both memory researchers and parapsychologists are missing an opportunity to be genuine scientific pioneers. Rather than boldly searching for new explanatory strategies (for memory specifically and for human behavior generally), they cling instead to familiar mechanistic presuppositions, which

they've typically never examined in any depth but by means of which they can maintain the illusion that they're doing science according to the allegedly tough-minded methods exemplified in some physical sciences. (Sherry Turkle has appropriately called this "physics envy."[32]) They can't get past the assumption that human abilities and behavior must be analyzed in terms of lower-level processes and mechanisms. And many seem not to recognize the difference between claiming that cognitive functions are *analyzable* in terms of underlying physical processes and claiming instead that those functions are merely *mediated* (perhaps only contingently) by underlying physical processes. But there are novel explanatory options and strategies they never consider; there are alternative and profoundly different approaches to the understanding of human beings. In chapter 3 I'll pursue that topic in more detail.

REFERENCES

Almeder, R. (1992). *Death and personal survival* . Lanham, MD: Rowman & Littlefield.

Beach, F. A., Hebb, D. O., Morgan, C. T., & Nissen, H. W. (Eds.). (1960). *The neuropsychology of Lashley: Selected papers of K. S. Lashley*. New York, NY: McGraw-Hill.

Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Oxford: Blackwell.

Braude, S. E. (1997). *The limits of influence: Psychokinesis and the philosophy of science* (Rev. ed.). Lanham, MD: University Press of America.

Braude, S. E. (2002). *ESP and psychokinesis: A philosophical examination* (Rev. ed.). Parkland, FL: Brown Walker.

Bursen, H. A. (1978). *Dismantling the memory machine : A philosophical investigation of machine theories of memory*. Dordrecht, Netherlands: D. Reidel.

Byrne, J. H. (Ed.). (2008). *Learning and memory: A comprehensive reference* (Vols. 1–4). Oxford: Academic.

Damasio, A. R. (1994). *Descartes' error: Emotion, reason and the human brain*. New York: Putnam.

Gazzaniga, M. S., Mangun, G. R., & Ivry, R. B. (1998). *Cognitive neuroscience: The biology of the mind*. New York: Norton.

Heil, J. (1978). "Traces of things past." *Philosophy of Science*, 45: 60-72.

Kelly, M. T., Crary, J. F., & Sacktor, T. C. (2007). Regulation of protein kinase Mζ synthesis by multiple kinases in long-term potentiation. *Journal of Neuroscience, 27*, 3439–3444.

Kesner, R. P. (2007). Neurobiological view of memory. In R. P. Kesner & J. L. Martinez Jr. (Eds.), *Neurobiology of learning and memory* (2nd ed.) (pp. 271–304). Burlington, MA: Academic.

Köhler, W. (1947). *Gestalt psychology*. New York: Liveright.

Köhler, W. (1969). *The task of gestalt psychology*. Princeton, NJ: Princeton University Press.

Lashley, K. S. (1929). *Brain mechanisms and intelligence*. Chicago: University of Chicago Press.

Lashley, K. S. (1950). In search of the engram. *Symposia of the Society for Experimental Biology, 4*, 454–482.

Lee, A. M., Kanter, B. R., Wang, D., Lim, J. P., Zou, M. E., Qiu, C., . . . Messing, R. O. (2013). PKMζ null mice show normal learning and memory. *Nature, 493*(7432), 416–419. doi:10.1038/nature11803

Liu, X., Ramirez, S., Pang, P. T., Puryear, C. B., Govindarajan, K. D., & Tonegawa, S. (2012). Optogenetic stimulation of a hippocampal engram activates fear memory recall. *Nature, 484*, 381–385.

Malcolm, N. (1977). *Memory and mind*. Ithaca, NY: Cornell University Press.
Menzel, R. (2008). Introduction and overview. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (Vols. 1–4) (Vol. 1, pp. 1–9). Oxford: Academic.
Moscovitch, M. (2000). Theories of memory and consciousness. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 609–626). Oxford: Oxford University Press.
Nadel, L. (2008). Multiple memory systems: A new view. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (Vols. 1–4) (Vol. 1, pp. 41–52). Oxford: Academic.
Nader, K. (2003). Memory traces unbound. *Trends in Neurosciences, 26*(2), 65–72.
Nader, K. (2007). A single standard for memory: The case for reconsolidation. *Debates in Neuroscience, 1*(1), 2–16.
Nader, K., & Einarsson, E. Ö. (2010). Memory reconsolidation—an update. *Annals of the New York Academy of Sciences, 1191*, 27–41.
O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Oxford University Press.
Pastalkova, E., Serrano, P., Pinkhasova, D., Wallace, E., Fenton, A. A., & Sacktor, T. C. (2006). Storage of spatial information by the maintenance mechanism of LTP. *Science, 313*, 1141–1144.
Pearsall, P., Schwartz, G. E. R., & Russek, L. G. S. (1999). Changes in heart transplant recipients that parallel the personalities of their donors. *Integrative Medicine, 2*(2/3), 65–72.
Pribram, K. H. (1971). *Languages of the brain*. Englewood Cliffs, NJ: Prentice-Hall.
Pribram, K. H. (1977). Holonomy and structure in the organization of perception. In U. M. Nicholas (Ed.), *Images, perception and knowledge*. Dordrecht, Netherlands: D. Reidel.
Pribram, K. H., Nuwer, M., & Baron, R. U. (1974). The holographic hypothesis of memory structure in brain function and perception. In D. H. Krantz, R. C. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vols. 1–2) (Vol. 2, pp. 415–467). San Francisco: Freeman.
Roll, W. G. (1983). The psi structure theory of survival. In W. G. Roll, J. Beloff, & R. White (Eds.), *Research in parapsychology 1982* (pp. 117–120). Metuchen, NJ: Scarecrow.
Sacktor, T. (2012). Memory maintenance by PKMζ—an evolutionary perspective. *Molecular Brain, 5*(1), 31.
Sacktor, T. C. (2011). How does PKMζ maintain long-term memory? *Nature Reviews Neuroscience, 12* (January), 9–15. doi:10.1038%2Fnrn2949
Schacter, D. L., & Tulving, E. (1994). *Memory systems*. Cambridge, MA: MIT Press.
Shema, R., Sacktor, T. C., & Dudai, Y. (2007). Rapid erasure of long-term memory associations in the cortex by an inhibitor of PKMζ. *Science, 317*, 951–953.
Stevenson, I. (1974). *Twenty cases suggestive of reincarnation* (2nd ed. rev.). Charlottesville: University Press of Virginia.
Tulving, E., & Craik, F. I. M. (Eds.). (2000). *Oxford handbook of memory*. Oxford: Oxford University Press.
Volk, L. J., Bachman, J. L., Johnson, R., Yu, Y., & Huganir, R. L. (2013). PKMζ is not required for hippocampal synaptic plasticity, learning and memory. *Nature, 493*(7432): 420–423. doi:10.1038/nature11802

NOTES

1. For representative samples of the view, see, e.g., Damasio, 1996; Gazzaniga, Mangun, & Ivry, 1998; Moscovitch, 2000; Tulving & Craik, 2000.
2. Similar problems also undermine theorizing in areas often related to parapsychology—for example, Sheldrake's account of morphic resonance. I deal with that in more detail in chapter 2
3. For extended critiques, see Bennett & Hacker, 2003; Braude, 2002; Bursen, 1978; Heil, 1981; Malcolm, 1977.
4. Heil, 1978 .
5. See, e.g., Köhler, 1947, 1969.
6. Köhler, 1969, p. 122, emphasis added.

7. Braude, 1997, 2002.
8. Nader, 2007, p. 2.
9. See also Nader, 2003; Nader & Einarsson, 2010.
10. See, e.g., Kelly, Crary, & Sacktor, 2007; Pastalkova et al., 2006; T. Sacktor, 2012; T. C. Sacktor, 2011; Shema, Sacktor, & Dudai, 2007.
11. T. C. Sacktor, 2011; emphasis added.
12. See, e.g., Lee et al., 2013; Volk, Bachman, Johnson, Yu, & Huganir, 2013).
13. Nader, 2003, 65.
14. Byrne, 2008.
15. Menzel, 2008, p. 6.
16. Ibid.
17. See, e.g., O'Keefe & Nadel, 1978; Schacter & Tulving, 1994).
18. O'Keefe & Nadel, 1978, p. 373.
19. O'Keefe & Nadel, 1978, p. 374.
20. Nadel, 2008, p. 45.
21. Kesner, 2007, p. 272.
22. Ibid.
23. Ibid.
24. Liu et al., 2012, p. 381.
25. Ibid.
26. Ibid.
27. Roll, 1983.
28. See, e.g., Almeder, 1992; Stevenson, 1974.
29. Pearsall, Schwartz, & Russek, 1999.
30. Beach, Hebb, Morgan, & Nissen, 1960; Lashley, 1929, 1950.
31. Pribram, 1971, 1977; Pribram, Nuwer, & Baron, 1974.
32. For more on that topic, see chapter 8.