CHAPTER 2

William James, Chaos Theory, and Conscious Experience

Andrew R. Bailey

Introduction

There is a plausible case—which as far as I know has not been made in print before—for the claim that, had William James had the terminology at the turn of the century, he would have described the brain as a chaotic system, and consciousness (including perceptual consciousness\(^1\)) as closely analogous to a strange attractor for that system. I want to make that case here.

There is some intrinsic interest, I think, in discovering that one of the founders of modern psychology, a figure “as close to being an American culture-hero as any academic...could be” (Wilshire, 1984, p. xvii), was a fledgling chaos theorist about the mind over one hundred years ago.\(^2\) Perhaps more importantly, however, James’ reasons for adopting a quasi-chaotic model of the brain still have a great deal of relevance today. Essentially, James saw his model as (so to speak) an attractive solution to the following problem: What are we to say about consciousness if, as he did, we hold all the following plausible propositions to be true together?

(a) Consciousness of some sort exists—it is a phenomenon for which we need a theory.
(b) Consciousness is not literally identical with or reducible to low-level descriptions of the physical brain.
(c) Consciousness is not a nonphysical entity (and therefore not a nonphysical causal agent).
(d) Consciousness has a causal role—it is not epiphenomenal.

These four propositions are each still widely attractive to contemporary workers in psychology, philosophy, and the brain sciences. Few thoughtful people today deny that some account of “consciousness” is needed for a complete science of the mind (though they may disagree strongly as to just what has to be explained; see, e.g., Marcel & Bisiach, 1988); and it is currently deeply unfashionable to believe in ghostly, Cartesian, mental stuff or objects (as opposed to peculiarly “mental” properties, which are still a live issue; see, e.g., Warner & Szubka, 1994). There is somewhat less consensus on Propositions (b) and (d); however, even most of those who deny (d), such as Chalmers (1996), admit that they would prefer not to do so if their theories did not push them towards it (and it
is generally accepted that epiphenomenalism raises difficult questions about, for example, the evolutionary role of consciousness; e.g., Heil & Mele, 1995; Sutherland, 1995). Regarding (b), it is usually taken as rote by at least the philosophers that mental properties cannot be identifiable with low-level physical properties (for several, quite complicated reasons, but mostly because, for any specific physical property—such as being in a certain configuration of neuronal firings—it is believed that entities in which that low-level property is not present can still have the mental property in question; see Block, 1980; Horgan 1984; Warner & Szubka, 1994). As a consequence, it is common throughout the study of the mind to speak of consciousness as being “emergent from,” or “at a higher level than,” or “supervenient upon,” the physical, by which it is typically meant that conscious mental states—such as the visual image of a red tomato—are somehow dependent upon or brought about by the physics of the brain, but need not be identifiable with any particular type of brain state (for example, there is no precisely specifiable “image-of-a-red-tomato”-type of brain state).3

There is, therefore, a wide range of central and influential thinkers, from Daniel Dennett (1991a, 1991b) to Petra Stoerig (1995) to Roger Penrose (1994), who would almost certainly agree with James and accept all of (a) to (d). However, as I think James realized (unlike several of those cited above), (a) to (d) together give the problem of theorizing about consciousness and the mind a particular, rather thorny, form. In essence:

1. What sort of thing can (perceptual) consciousness possibly be if it is neither a separate “mental” entity nor identical with a physical thing?

2. How can (perceptual) consciousness be causally efficacious if (i) it is neither a cause nor an effect, as we usually understand them—if, since it is not an entity, it is apparently not the sort of thing that can have causal interactions? or (ii) only low-level physical properties can have causal powers—if, for example, the universe is causally closed at the level of micro-physics, a level at which, from Proposition (b) above, the property of experiential consciousness is not present?

These are the questions that James’ chaotic theory of the mind is formulated to answer. Naturally, the various general theories of the mind and consciousness put forward by the diverse thinkers cited above (and many others) have cast several different shades of light upon these problems, and have sometimes even addressed them directly. However, none of them have met with anything close to universal acceptance. The chaotic model found in James, therefore, provides an alternative solution which might fruitfully be explored in the context of the modern debate.

In what follows I shall first quickly describe what I take to be “chaos theory” so that it can be compared to the relevant aspects of William James’ account of consciousness, which I will develop through the following three sections. I will textually demonstrate James’ espousal of the four points (a) to (d) above, and show how considerations about the causal role of consciousness in particular led him to put forward a model in which (implicitly) the brain is a chaotic system and consciousness its attractor. In section four I hope to make
plausible the claim that James did indeed present an account which is strikingly similar to current chaos-theoretical models in other domains. Finally, I will very briefly make some points about the general plausibility, and potential fruitfulness, of a chaotic model of consciousness and the brain. Following James, I shall usually speak quite generally of “consciousness” and “the brain”: The processes of perception in particular should simply form a specific case—indeed, quite a central case—for these general claims.

A Definition of “Chaos Theory”

Strictly speaking there is no such thing as “chaos theory.” Or rather, most scientists and mathematicians rarely use the expression in their formal work, and refer only to investigations of “chaotic phenomena” or “dynamical chaos” (though what exactly makes a system “chaotic” is also imprecise, unless it is just that it is describable using the tools of chaos theory). The guiding insight behind the notion of chaos is that systems governed by mathematically simple equations can exhibit elaborately complex, de facto unpredictable, behavior. This unpredictability results from “sensitive dependence on initial conditions,” which means that very—indeed, vanishingly-small—measurement inexactness will become tremendously magnified in a short time.

Arguably one of the most precise definitions of chaos theory was formulated by Stephen Kellert in 1993, as follows: Chaos theory is “the qualitative study of unstable aperiodic behavior in deterministic nonlinear dynamical systems” (Kellert, 1993, p. 2). A dynamical system is a numerical model—encompassing the changing values of some selective set of variables—of an actual segment of the world, which contains “evolution equations” for the calculation of changes in state of that system. For chaotic systems the evolution equations always include nonlinear terms, which makes “closed-form” solutions of these equations impossible—roughly, a closed-form solution is a single formula that allows one to simply plug in the time of the desired prediction into the equation and determine the state of the system at that time. By contrast, “open-form” solutions require one to recalculate the state of the system for each temporal increment between the initial time and the final time; this is such a computationally unwieldy method that typically researchers in chaotic domains focus, not upon making exact predictions of future states, but upon using mathematical techniques to “provide some idea about the long-term behavior of the solutions” (Devaney, 1986, p. 4). This is what is meant here by “qualitative study” (which is not to say, of course, that it is not technical or mathematical in character).

The behavior of a system is unstable if it never settles into a form that resists small disturbances—unlike, say, a marble rolling into the bottom of a bowl—and is aperiodic if no variable describing the system undergoes a regular repetition of values. Unstable periodic behavior, then, is very complex: It never repeats and continues to manifest the effects of any small perturbation. This
produces a series of measurements that appear absolutely random: One standard example of this kind of behavior is the course of human history. Finally, however, unlike human history, chaotic systems are called “deterministic” because they are composed of only a few—typically, fewer than five—differential or difference equations, which make no explicit reference to chance mechanisms. Central examples are theories of fluid turbulence [though, as David Ruelle (1991) has noted, “the subject does not seem to have revealed its last mysteries” (p. 52)] or population dynamics.

A chaotic system, then, is one which is describable in the terms of chaos theory: roughly, one whose behavior is unstable and aperiodic and which can be modeled using deterministic nonlinear equations.

We shall also need to make use of the notion of a “strange attractor.” An attractor is an object with no volume in state space towards which all nearby trajectories will converge. (State space is a mathematically constructed conceptual space in which each dimension corresponds to one variable of the system—thus, each point in state space represents a full description of one possible state of the system, and a “trajectory” through state space represents an evolution of the system through time.) Attractors can be generated by the numerical integration of a simple set of dynamical equations; however attractors are emergent phenomena in dynamical systems and it is very hard indeed to tell what kind of attractor is present by looking at the equations.

Motion on a strange attractor will exhibit sensitive dependence on initial conditions: That is, for any point on the attractor, there is another point nearby that will follow a path diverging exponentially from the path of the first. The apparent tension between the attraction of an attractor—pulling trajectories together—and the strangeness of its sensitivity to initial conditions—pushing trajectories apart—underlies the striking geometric feature possessed by strange attractors of both “stretching” and “folding.” The action of a chaotic system takes nearby points and stretches them apart in a certain direction, creating the local divergence responsible for unpredictability; on the other hand, the system also “folds” together distant points, causing trajectories to converge in a different direction. Thus, nearby points can quickly evolve to opposite sides of the attractor, yet the trajectories are confined to a particularly shaped region of phase space.7

Because of this stretching and folding, strange attractors typically have fractal shapes—that is, they can usually be treated as a stack of two-dimensional sheets displaying a self-similar packing structure—and so have a nonintegral number of dimensions.8

A strange attractor, therefore, is an emergent object with no volume in state space towards which all nearby trajectories will converge and the (fractal) shape of which exhibits sensitive dependence upon initial conditions.
James on the Insubstantiality of Consciousness: The Ontological Difficulty

Despite having written a paper entitled *Does “consciousness” exist?* (1904) which famously answered the question in the negative, it is nevertheless clear that William James did, throughout his lifetime, assert the existence of some sort of consciousness. As he put it in his *Psychology: Briefer Course*, “the first and foremost concrete fact which everyone will affirm to belong to his inner experience is the fact that consciousness of some sort goes on” (1892, p. 140), and he never deviated from this early view. What James was concerned to combat was a particular bundle of views about consciousness—roughly those which treat consciousness as some sort of object, or series of objects, either physical or “spiritual.” Thus:

I mean only to deny that the word stands for an entity, but to insist most emphatically that it does stand for a function. There is, I mean, no aboriginal stuff or quality of being, contrasted with that of which material objects are made, out of which our thoughts of them are made; but there is a function in experience which thoughts perform, and for the performance of which this quality of being is invoked. (1904, p. 4)

Let us begin, then, by outlining what James did not want to dispute about consciousness—that is, in roughly what form he held Proposition (a), that consciousness does exist, to be true. Essentially, what James felt was a set of phenomena in need of a theory was no more and no less than the stream of consciousness. In his preface to the *Principles of Psychology* (1890) James explains that he will restrict himself to what he calls the data of consciousness—that is, “thoughts and feelings”—and put aside any “attempts to explain our phenomenally given thoughts as products of deeper-lying entities” as being “metaphysical” (1890, Vol. 1, p. vi). Unlike the “metaphysics” of mind, James held, our thoughts and feelings are simply undeniable: when we introspect “everyone agrees that we there discover states of consciousness. So far as I know the existence of such states has never been doubted by any critic, however sceptical in other respects he may have been.... I regard this belief as the most fundamental of all the postulates in Psychology...” (1890, Vol. 1, p. 185). So, states of consciousness indubitably exist and “[b]y states of consciousness are meant such things as sensations, desires, emotions, cognitions, reasonings, decisions, volitions, and the like” (1892, p. 9). Even in 1904–05, whilst in the throes of attacking the very notion of “consciousness,” James felt able to assert that “thoughts in the concrete are fully real” (1904, p. 19), and that, even after the total demolition of the ontological dualism of consciousness, “that which we
suppose to exist is...that which until now has been called the content, the Inhalt, of consciousness” (1905, p. 7).

However, James is at some pains to point out, the indubitable data of consciousness do not license any strong metaphysical conclusions about the “stuff” of consciousness. All that we can conclude from careful empirical examination of the evidence is that “thought goes on” (1890, Vol. 1, p. 225). More specifically, the limit of our certain introspective acquaintance is a sequence of thoughts with the following five famous characteristics (1890, chap. 9):

- every thought is part of a personal consciousness;
- thought is in constant change;
- within each personal consciousness thought is sensibly continuous;
- thought appears to deal with objects independent of itself; and
- thought is selectively attentive.

Nothing more can be said about consciousness that is not mere ungrounded metaphysical speculation. And, James later wrote, “I believe that “consciousness,” when once it has evaporated to this estate of pure diaphaneity, ...is on the point of disappearing altogether. It is the name of a nonentity” (1904, p. 3). James was quite explicit throughout his writings—especially his later writings—that:

1. Consciousness is not like a container or a medium—it, itself, is just the continuous stream of thought. Thus James inveighs against the contemporary neo-Kantians—“belated drinkers at the Kantian spring”—who supposed that we “have an immediate consciousness of consciousness itself... [which] is believed to ... be felt as a kind of impalpable inner flowing” (1904, p. 5). And he disavows the view—commonly held by psychologists of his time—that consciousness is like the pure solvent or menstruum of paint which, only when mixed with the pigment of thoughts and feelings, can make up the rich colors of mental life (1904, p. 6): Instead, James holds, experience is wholly made up of thoughts and feelings, and nothing else.

2. There is no empirical, introspective evidence for a soul, ego, or knower behind or within our consciousness. Rather, James insisted, consciousness is “natively selective,” not “directed” by some other agency distinct from the stream of thought itself. There is no entity—“consciousness”—juxtaposed to the thoughts themselves, the “contents” which make up the stream of thought.

[E]xperience, I believe, has no such inner duplicity; and the separation of it into consciousness and content comes, not by way of subtraction, but by way of addition—the addition, to a given concrete piece of it, of other sets of experiences.... [A] given undivided portion of experience, taken in one context of associates, play[s] the part of a knower, of a state of mind, of ‘consciousness’; while in a different context the same undivided bit of experience
plays the part of a thing known, of an objective ‘content.’” (1904, pp. 6–7)

3. There is no evidence that the stream of consciousness is made up of some different kind of substance than the “rest” of the world. James unequivocally rejects the view of our mental life “as a sort of interior current—active, light, fluid, delicate, diaphanous, so to speak—and absolutely different from what is material” (1905, p. 2). Indeed, he goes so far as to say unqualifiedly that “thoughts in the concrete are made of the same stuff as things are” (1904, p. 19).

I believe that consciousness (as it is commonly represented, either as an entity, or as pure activity, but in any case as being fluid, unextended, diaphanous, devoid of content of its own, but directly self-knowing-spiritual, in short), I believe, I say, that this sort of consciousness is pure fancy. (1905, p. 7)

In sum, then, it seems perfectly clear that throughout his major writings James endorsed Proposition (c) above: that, whatever else it may be, consciousness is certainly not some kind of nonphysical substance, process or entity.

Is consciousness, then, a physical entity for James? Is it, say, just a certain class of the operations of the brain? It is not. First, we have already seen that James does not think of consciousness as an entity in juxtaposition or addition to the stream of thought at all—although, because of the historical context in which he was writing, James attacked the notion of extra “soul-stuff,” he could equally well have reacted against the notion of, say, a kind of physical “self” or “mind’s eye” in the brain, or some more physicalist view of a “stream” in which the contents of thought bob along. Second, James frequently denies that the thoughts and feelings themselves—the components of the evidently present stream of consciousness—are to be identified with states or activities of the brain—instead, James restricts himself to the claim that they are merely correlated with activities of the brain.

The consciousness, which is itself an integral thing not made of parts, “corresponds” to the entire activity of the brain, whatever that may be, at the moment. This is the way of expressing the relation of mind and brain from which I shall not depart during the remainder of the book, because it expresses the bare phenomenal fact with no hypothesis, and is exposed to no such logical objections as we have found to cling to the theory of ideas in combination. (1890, Vol. 1, p. 177)
Mental state-types and property-types (e.g., a visual experience of redness) are, for James, not the same thing as types of physical state or property (e.g., a certain kind of agitation in the occipital lobe)—this is Proposition (b) above.

Although we affirm that the coming to pass of thought is a consequence of mechanical laws...we do not in the least explain the nature of thought by affirming this dependence, and in that latter sense our proposition is not materialism. (1890, p. 13)

The “admitted fact of functional dependence”—what James once called “the great psycho-physiological formula: Thought is a function of the brain” (1897, p. 10)—is for him always a relation between separate terms. Indeed, when James lists the possible varieties of functional dependence—productive, permissive, or transmissive—none of them constitute self-identity (1897, pp. 12–14).14

Further, James lists many attributes of consciousness which he considers not to be predicable of the physical. For example, the quotation from page 177 of the Principles above shows that James held consciousness to be indivisible into separate parts—he devotes the bulk of Chapter VI of Volume 1 of the Principles to arguing this point—yet the brain is clearly built up of atomic parts: How then can the two be the same thing? He also dwells at some length on the fact that consciousness is unique in that it can have “ends” or “interests.” Considered merely physically, by contrast, the reactions of our brain, cannot be properly talked of as “useful” or “hurtful” at all.... All that can be said of them is that if they occur in a certain way survival will as a matter of fact be their incidental consequence. The organs themselves, and the rest of the physical world, will, however, all the time be quite indifferent to this consequence, and would quite as cheerfully, if the circumstances changed, compass the animal’s destruction. (1890, p. 141)

Further,

Nothing can more strikingly show, it seems to me, the essential difference between the point of view of consciousness and that of outward existence. We can describe the latter only in teleological terms, hypothetically, or else by the addition of a supposed contemplating mind which measures what it sees going on by its private teleological standard, and judges it intelligent. But consciousness itself is not merely intelligent in this sense. It is intelligent intelligence. It seems both to supply the means and the standard by which they are measured. It not only serves a final
purpose, but *brings* a final purpose—posits, declares it. (1878, pp. 27–28)

Other examples of nonphysical predicates of thoughts and feelings, adduced by James, are the property of “knowing” or “reporting” and of being “personal.” In addition, all sorts of things are true of “mental objects” (that fire may play over them and not affect them, that they only came into existence moments ago) that are false of their corresponding “physical” counterparts, and *vice versa* (1904, pp. 8–9).

James, then, held that consciousness—in the form of the stream of thought—*certainly exists*, but is neither some kind of nonphysical stuff or process *nor*, though it is closely linked with the brain, is it *identifiable* with aspects of the brain. This brings James’ metaphysical quandary to a sharp point. What *are* thoughts and feelings, if they are neither bits of spirit nor bits of the brain (nor, presumably, extra bits of material stuff floating around in addition to the brain)? What exactly, for a Jamesian, *is* “the stuff that dreams are made on”?

**James on Interactive Consciousness: The Causal Difficulty**

It is quite clear from James’ writings, in particular Chapter 5 of his *Principles*, that he was firmly opposed to epiphenomenalism—to any view which “banishes [consciousness] to a limbo of causal inertness” (1890, Vol. 1, p. 135), where it exists more like a “melody,” or a “shadow” than like a “real thing.” As he states firmly, “[i]t is to my mind quite inconceivable that consciousness should have *nothing to do* with a business which it so faithfully attends” (1890, Vol. 1, p. 136).

He argues explicitly against various reasons one might suppose consciousness to be causally inert (1890, Vol. 1, pp. 133–138), and then proceeds to put forward a positive argument to the effect that “the particulars of the distribution of consciousness, so far as we know them, point to its being efficacious” (1890, Vol. 1, p. 138):

(i) “Consciousness grows the more complex and intense the higher we rise in the animal kingdom” (1890, Vol. 1, p. 138);

(ii) Consciousness might help “maintain the animal in the struggle for existence” (1890, Vol. 1, p. 138);

(iii) The defects of the “other” human organs “are such as to make them need just the kind of help that consciousness would bring provided it were efficacious” (1890, Vol. 1, p. 138); and

(iv) Therefore “the plausible inference [is] that it came just because of its efficacy” (1890, Vol. 1, pp. 138–139).

Evidently, then, James endorsed our Proposition (d) above: that consciousness is causally efficacious. One’s conscious states are causally relevant—they *make a difference* to our subsequent behavior. However this, in combination with his endorsement of (a) to (c), immediately raises a second
serious difficulty for James’ account of the mind: It is not immediately clear how James can consistently hold consciousness to be causal.

As we have seen, James is a parallelist in the sense that for him, states of consciousness are correlated with, but distinct from, states of the brain. Further, James certainly held that the operations of the brain are causal. Then, as James was very well aware, it is a small, shuffling step from here to the position that “even where we know consciousness to be there, the still more complicated neural action which we believe to be its inseparable is alone and of itself the real agent of whatever intelligent deeds may appear” (1890, Vol. 1, p. 129). In other words, it is natural to think of the brain as a causally complete system, fully capable in principle of explaining all of our behavior; the states of consciousness with which these brain processes happen to be correlated play no role whatsoever in bringing about that behavior (which conceivably could occur in just the same way even in the complete absence of consciousness). Not only would this mean that mental life has no effect on the physical, but nor would mental events even cause each other, since feelings would not cause nerve-actions: Any mental state would be “only the correlate of some nerve-movement whose cause lay wholly in a previous nerve-movement” (1890, Vol. 1, p. 133). Thus, as one critic puts it,

[o]n the one hand James tells us that consciousness directs thought and action, while on the other hand he pictures consciousness as only parallel to the bodily nervous processes that are thought and action.

(Leahey, 1980, p. 266)

Though James clearly rejects this epiphenomenalism—or “automaton-theory” as he calls it—how can he justify this denial? One natural tack would be to suggest that mental states are additional causes of behavior—that the physical brain causally underdetermines some of the behavior of a conscious organism, and that this causal gap is filled by consciousness. As he suggests at one point, “mental life seems to intervene between impressions made from without upon the body, and reactions of the body upon the outer world again” (1890, Vol. 1, p. 6); it would solve James’ causal problem if he could suppose that thoughts and feelings are intermediate links in this causal chain. Perhaps, a simple story could run, afferent stimulus A (brought about by a candle flame, maybe) causes a sequence of impulses which bring about mental state M (a sensation of pain) which then in turn precipitates a stream of efferent impulses ending in behavior B (snatching back one’s hand and uttering a cry, say); without M, then, the causal chain would be incomplete and, presumably, no behavior (or possibly merely reflex behavior) would result.

However, this kind of response to the epiphenomenal challenge is simply not available to James. There is nothing in his account which can possibly fill the place of M. As we have already seen, thoughts and feelings, for James, are not little bits of stuff; they are not like ectoplasmic billiard balls, capable of receiving a causal impulse and transferring it to some other, perhaps physical, entity. There
are no entities around at all, as far as James’ view of consciousness is concerned; hence, it is very hard to see how there could be causal agents. The following, then, is distinguished James scholar Gerald Myers’ (1986) final verdict:

Whether the phenomenon is called experience or consciousness, James’ testimonial words to it lead us irresistibly to view it not merely as a function but as something with an inherent nature, by which it is a causal agent that produces effects. James wanted to hold that in one way consciousness does not exist, but that in another way it does; yet he was never able, even to his own satisfaction, to define the two ways clearly enough to show that they are consistent rather than contradictory. (p. 64)

**Consciousness and Chaos: A Solution**

James’ account of consciousness, then, apparently faces two connected problems: It is ontologically vague, and it lacks an account of the causal efficacy of conscious states. Further, these problems are shared by any account of the mind that also endorses all four of Propositions (a) to (d). Let us make this a little more concrete for the domain of perception in particular. The Jamesian view holds that, during perception, there are two processes occurring in parallel: a sequence of physical changes to the brain, and a (shorter) series of correlated and dependent, but non-identical, mental states, such as visual images, smells, and sounds. These mental states are neither chunks of “spirit” nor are they bits of the brain; they do not inhere in any kind of substantial medium; they do not constitute any entity like a “self” or “ego.” What then, metaphysically speaking, are they? Despite their insubstantiality, these conscious states are causally relevant to perception: They make a difference to the kinds of behavior that are stimulated by perceptual states. Roughly speaking, perhaps, for the Jamesian one would not normally back away from a snake coiled to strike unless one had both a visual image of that snake and a sensation of fear; one would not go to the record store and buy CDs unless one experienced pleasurable sensations in listening to music; and so on. Exactly how, in that case, do conscious sensations exert a causal influence, over and above the deterministic machinations of the brain itself?

James’ solution, I submit, was to suggest (as best he was able using the conceptual tools of his time) that consciousness is best treated as an attractor for a chaotic system, namely, the brain (or, at least, the cerebral cortex). This is what he means, in the third premise of the argument for conscious causation outlined above, when he says that “the defects of [the]...other organs (where consciousness is most developed) are such as to make them need just the kind of help that consciousness would bring provided it were efficacious” (1890, Vol. 1, p. 138).

But what are now the defects of the nervous system in those animals whose consciousness seems most highly developed? Chief among
them must be instability. The cerebral hemispheres are the characteristically “high” nerve-centres, and we saw how indeterminate and unforeseeable their performances were in comparison with those of the basal ganglia and the cord. But this very vagueness constitutes their advantage. They allow their possessor to adapt his conduct to the minutest alterations of the environing circumstances. (1890, Vol. 1, p. 139)

James goes on to even more explicitly attribute to the cerebral hemispheres characteristics which, as we showed above, are hallmarks of chaotic systems. We said that chaotic systems are those whose behavior is unstable, aperiodic, deterministic, and capable of being modeled using nonlinear equations. Not surprisingly, James did not address the last point; but he is quite clear on the preceding three. “An organ swayed by slight impression is an organ whose natural state is one of unstable equilibrium,” he writes.

We may imagine the various lines of discharge in the cerebrum to be almost on a par in point of permeability—what discharge a given small impression will produce may be called accidental, in the sense in which we say it is a matter of accident whether a rain-drop falling on a mountain ridge descend the eastern or western slope. (1890, Vol. 1, p. 139)

The movements of the brain per se yield the means of attaining these [proper] ends mechanically, but only out of a lot of other ends, if so they can be called, which are not the proper ones of the animal, but often quite opposed. The brain is an instrument of possibilities, but of no certainties. (1890, Vol. 1, p. 141)
What this means is that,

I do not see how one could reasonably expect from it [the cerebrum] any certain pursuance of useful lines of reaction.... We can never be sure that its equilibrium will be upset in the appropriate direction. In short, a high brain may do many things, and may do each of them at a very slight hint. But its hair-trigger organization makes of it a happy-go-lucky, hit-or-miss affair. It is as likely to do the crazy as the sane thing at any given moment.... The performances of a high brain are like dice thrown forever on a table. Unless they are loaded, what chance is there that the highest number will turn up oftener than the lowest? (1890, Vol. 1, p. 140)

Faced with a sheer cliff, the happy-go-lucky brain may cause its organism to either retreat or leap off, depending on such incidentals as the prevailing wind-speed or the agent’s favorite color.

Of course, human beings do not actually respond randomly to their environment, and they are on the whole substantially more likely to behave sanely than madly. James, naturally, is perfectly well aware of this: “All this is said of the brain as a physical machine pure and simple” (1890, Vol. 1, p. 140). What makes the difference? It is consciousness which “loads the brain’s dice,” and causes us to behave in a sensible and goal-directed manner. It is consciousness which is the missing piece of the puzzle and which constrains the happy-go-lucky brain.

How does consciousness perform this task? It does so, James says, by “bringing a more or less constant pressure to bear in favor of those of its performances which make for the most permanent interests of the brain’s owner.” It exercises “a constant inhibition of the tendencies to stray aside” (1890, Vol. 1, p. 140).

Feeling, from this point of view, may be likened to a cross-section of the chain of nervous discharge, ascertaining the links already laid down, and groping for the one which seems best to fit the case. (1890, Vol. 1, p. 142)

That is, perhaps we may very loosely paraphrase, from a range of possible paths in state space, consciousness “pulls” the behavior of the brain in a certain direction.

Significantly, consciousness, for James, does not initiate new actions, but instead it modifies—reinforces and inhibits—the complex, hair-trigger reflex actions of the brain.

The feelings can produce absolutely nothing new, they can only reinforce and inhibit reflex currents which already exist, and the
original organization of these by physiological forces must always be the ground-work of the physiological scheme. (1890, Vol. 1, p. 138)

Perhaps because he does not have the vocabulary of attractors, James tends to speak of consciousness in this context as some kind of teleological agent superadded to the physical brain: as a “fighter for ends” which is “primarily a selecting agency” (e.g., 1890, Vol. 1, p. 139). However, as we have seen, and as his critics have been quick to point out, this kind of language is strictly speaking unavailable to James—this cannot literally be what James meant. How suggestive, then, are James’ comments of the anachronistic notion of consciousness being a strange attractor? Sufficiently so, I think, that we might say that, even if this was not the concept he was groping towards, it at least is one which slots very nicely into the oddly-shaped theoretical space he has left for consciousness.

A strange attractor, we said above, is an emergent object with no volume in state space towards which all nearby trajectories will converge, and the (fractal) shape of which exhibits sensitive dependence upon initial conditions. It is what we might call an emergent structural feature of chaotic systems, but has no ontological existence in addition to those systems; nor, of course, is it identifiable with any of the states or parts of that system. This squares very nicely indeed with what James has to say about consciousness.

Causally, for James, consciousness is not some kind of substantial causal agent yet it restricts the chaotic behavior of the brain—it “brings it about” that the behavior of the system remains within certain “intelligent” parameters, and does not behave “crazily” or self-destructively. That is, in the absence of consciousness (an attractor), the brain would flail absolutely wildly in response to environmental stimuli. When consciousness (an attractor) is present, the system still behaves unpredictably in detail (James would surely agree that human behavior is like this) but typically remains within certain parameters, and over time exhibits some kind of complex pattern.

Ontologically, consciousness—if it is an attractor—is not an “entity,” “process,” or “stuff,” but nevertheless is perfectly real as something like a “principle of functional organization” of the system as a whole. It is neither made of physical nor nonphysical stuff, yet is completely dependent upon the physical system of which it is a pattern—if the system (the brain) did not exist, then neither would the attractor (consciousness). Finally, the nature of consciousness is not evident from mere inspection of the brain and its laws, just as the shape of an attractor can typically only be found by repeated measurements of the system over time.

Adopting the view that consciousness is a strange attractor for the chaotic system which is the cerebrum, then, is not only consistent with at least a large swathe of textual evidence from James but apparently also solves both of the problems with his account of the mind (and other similar accounts) which we
have identified here. Further, it is difficult to see how James could solve these problems of internal consistency without endorsing a view of this sort.

**Conclusion: Prospects for the Chaotic Model**

I have argued that William James presented a view of consciousness and the brain which comes very close indeed to being the theory that the higher brain is a chaotic system and consciousness its strange attractor. In addition, I have claimed that this thesis is a response to dilemmas raised by holding four plausible theses about the mind in conjunction; that these four theses are quite widely endorsed; and that therefore this chaotic model of consciousness, including perceptual consciousness, is a live option—possibly even a necessary resort—for many present day workers in the field.

It seems a shame to close this essay, however, without any speculations about just how fruitful this chaotic model of the mind might turn out to be. Personally, I am not entirely sanguine about its prospects, for a number of reasons. First, even given the truth of everything James has to say on the subject, it is still not entirely clear that the brain really is, strictly speaking, a chaotic system; this is because, though as a whole it shows sensitive dependence upon initial conditions, it is far from certain that its activities can be modeled by a small number of nonlinear equations. That is, it remains to be shown that the cerebral hemispheres possess the hallmark of chaos: producing a large degree of complexity from a system which can be modeled mathematically simply.

Second, though the equation of consciousness with an attractor has some ontological attractiveness, from the perspective of the Jamesian, it seems to me to lack introspective plausibility. An attractor is merely a topological object in phase space, and has no more reality than this; by contrast, visual sensations of redness, experiences of intense pain, the taste of a lemon, and the smell of rotten eggs have a much more forceful claim to ontological existence (even given all of James’ positivism). Indeed, how could an attractor—a mathematical object—be a phenomenal flux as James describes it, and as we experience it?

Third, the attractor solution to the causal problem, at least in the basic form in which I have presented it here, lacks sufficient specificity: All that is de facto predictable about a chaotic system, typically, is that its successive states will appear somewhere on the attractor—just which succession of states will occur remains mysterious. By contrast, with human beings we are not just limited to predicting that their behavior will be somewhere (anywhere) within the bounds of what we might call “reasonable”: We can narrow down our predictions much more accurately than that. In general, that is, we can not only predict that our acquaintances will not act wildly; we can predict that they will generally arrive at certain places when they say they will, will carry umbrellas when it rains, will fill up their car with gas when the fuel gauge is near the empty mark, and so on.

My provisional response to these difficulties would be to call into question the necessity for holding all four of the Propositions (a) to (d). That is, it might be
that holding (a) to (d) commits one to a chaotic model of consciousness (though I have not argued this very strong claim here); further, perhaps the chaotic model will ultimately turn out to be unfruitful or implausible. In that case, the only reasonable response would be to jettison one of the four assumptions—probably Proposition (b), which forbids the straightforward identification of mental states with physical states. Without this assumption, we can say that mental state-types—such as the visual sensation of redness or the experienced scent of rotten eggs—are identical with physical state-types, and as such have a clear ontological status (being physical) and causal role (since they are the same thing as certain, causally influential, physical properties). Unfortunately, however, to abrogate Proposition (b) would be simultaneously to reject large parts of William James’ stimulating and filigreed account of the mind.

Notes

1By “perceptual consciousness” I mean all the experientially conscious aspects of perception—any aspect of perception that there is “something it is like” to undergo. This includes visual images, auditory experiences, tastes, etc., as well as, perhaps, the deliberate judgment that something is P or the conscious decision to, say, attend more fully in a certain direction. The question discussed in this essay, then, in part, deals with the role that perceptual consciousness has within the whole (conscious and unconscious) process of perception.

2Perhaps unfortunately, there is a bit of a bandwagon phenomenon here: It has become fashionable to link one’s discipline or subject area with chaos theory. Not only have the tools of chaos theory been applied with varying success to domains ranging from economics and business (Nilson, 1995; Trippi, 1995) to history (Beaumont, 1994) and even criminology (Milovanovic, 1997), but its ideas have been said to parallel and confirm literary theory (Hawkins, 1995; Hayles, 1990; Livingston, 1997), pop psychology (Wieland-Burston, 1992) and Taoism (Briggs & Peat, 1989; Walter, 1996), while Baudrillard has recently proposed a fractal model of the postmodern self. Personally, I hope here to avoid the trap of merely linking two currently modish domains.

3For myself, I find Proposition (b) conceptually muddy and probably false; further, the problems of causation and ontology with which this paper deals mostly fall away if this assumption is removed. However, there is no doubt that (b) is very widely accepted today, so a plausible account of the mind which allowed it to stand would be of great interest—this is the line I pursue here.

4Apparently the term “chaos” was first coined by Jim Yorke, an applied mathematician from the University of Maryland, in 1975 (Li & Yorke, 1975). However, in fact, the kind of time evolution with many periodic orbits described
in this paper turned out not to be a phenomenon with sensitive dependence on initial conditions at all!

5That is, terms which involve algebraic or other more complicated functions of the system variables (e.g., \(x^2\) or \(\sin(x)\) or \(5xy\)).

6Typical questions in this kind of work might be: What characteristics will all solutions of this system ultimately exhibit? Or: How does this system change from exhibiting one kind of behavior to another kind?

7This has been compared to the operation of a pinball machine: The pins “stretch” the trajectory of the ball, but the edges of the table keep it “folded.” “What do you do if you’re not allowed to behave simply, but you can’t get away? You are forced to do something complicated. The pinball bounces from pin to pin, never doing the same thing twice” (Cohen & Stewart, 1994, p. 190).

8For example the Lorenz attractor must have fewer dimensions than three—or it would have volume in three dimensional space—but must also have more dimensions than two—otherwise trajectories of the system would have to cross and the system would no longer be deterministic...at a cross-point, the system would have a “choice” which way to go. Similarly, the Hénon attractor has a dimension between one and two.

9In fact, James perceptively (if modestly) commented that “in this strictly positivistic point of view consists the only feature of [The Principles of Psychology] for which I feel tempted to claim originality” (1890, Vol. 1, p. vi).

10As Alfred Schuetz puts it, James rejects “Locke’s simile of a ‘white paper’ soul” (Schuetz, 1941, p. 444).

11It must be admitted that James was much less dogmatic in his attack on the “soul” than he was on the general notion of a ghostly knower or self behind the stream of consciousness. For example, he says at one point, “our reasonings have not established the non-existence of the Soul; they have only proved its superfluity for scientific purposes.” I take it that James pulled his punches here because of the religious implications of eliminating the soul.

12The careful reader might object at this point that I am forgetting James’ well-known doctrine of radical empiricism, by which he replaces both mind-body dualism and materialist monism with the view that everything in the universe is made up of units of what he called “pure experience.” For reasons of space and focus I will not explicitly address James’ radical empiricism here. Note, however, that Proposition (c) is, roughly, the claim that traditional substance dualism is false—that there is no Cartesian “soul-stuff” with properties very different from
those of the material of the physical world. This is consistent with James’ monism, and is precisely the kind of position he attempts to refute in his writings on consciousness.

Some commentators—such as Myers (1986)—argue that James’ focus upon correlation rather than identification, especially in the *Principles*, was merely a presentational device, designed to avoid alienating his readership by the espousal of “materialism,” which was James’ real position. I do not believe that the textual evidence supports this interpretation.

When he does, at one point, note the possibility of self-identity, he (a little dismissively) calls it “the monistic theory,” where mind and brain are “inner and outer aspects” of “One and the Same Reality.” After mentioning the theory, he then proceeds to discuss its two alternatives (“the spiritualistic theory” and “the atomistic theory”) in more detail, suggesting that these are the two really live options (1892, p. 396).

(a) The argument from continuity—that basic brain functions, like for example the frog’s spinal cord, operate merely by “reflex” and that therefore all brain functions must do the same—is inadequate because exactly the same argument could be made in reverse: The higher operations of the brain are conscious, so all the operations of the brain must really be due to the, sometimes invisible, presence of a consciousness lower in degree (1890, Vol. 1, p. 134).

(b) It may be true that epiphenomenalism provides a way of “keeping things simple” for scientists of the mind by restricting theory to physical language, but such a move is illegitimate, in James’ view—it is precisely the role of psychology to explain the relationship between mind and brain (1890, Vol. 1, p. 136).

(c) It is also true that the causal relation between ideas and the brain is wholly mysterious—even inconceivable. But, James argues, the same is true of every instance of causation, as Hume is supposed to have demonstrated. “However inadequate our ideas of causal efficacy may be, we are less wide of the mark when we say that our ideas and feelings have it, than the Autonomists are when they say they haven’t it” (1890, Vol. 1, p. 137). Psychology, as a natural science has a duty to be naïf about the status of causes—if something seems like a cause, then they had better treat it as if it is (1890, Vol. 1, pp. 137–138).

“If we knew thoroughly the nervous system of Shakespeare, and as thoroughly all his enironing conditions, we should be able to show why at a certain period in his life his hand came to trace on certain sheets of paper those crabbed little black marks which we for shortness’ sake call the manuscript of Hamlet. We should understand the rationale for every erasure and alteration therein, and we should understand all this without in the slightest degree acknowledging the existence of the thoughts in Shakespeare’s mind” (1890, Vol.
1, p. 132). In current parlance, this is the view that “zombies”—creatures whose behavior is outwardly completely identical with a human being’s, but who experience no glimmerings of experiential consciousness—are logically possible. See Sutherland 1995.

17It is entertaining to recall that Jeff Goldblum’s chaos mathematician character in the film Jurassic Park exemplifies the phenomenon of chaos with a strikingly similar example—by showing that a drop of water, landing in the same place, will run in a different direction off someone’s hand each time because of tiny variances in initial conditions.

18On the other hand, it should be admitted, one could imagine ways of talking by which the attractor for one’s own brain might be said to have the five characteristics that James held are all that is evident about consciousness: being part of a personal consciousness, being in constant change, being sensibly continuous, and so on. I doubt James himself would be happy with this kind of sophistry though.

References


