

Phil. Soc. Sci. 12 (1982) 153-171

Julian Jaynes and the Bicameral Mind: A Case Study in the Sociology of Belief¹

W. T. JONES, *Division of the Humanities and Social Sciences, California Institute
of Technology*

That reasonable people are able to weigh evidence objectively and so reach agreement about matters of fact was an article of faith in the Age of Reason. It is still a presupposition of much thinking about thought. Thus Stephen Pepper's definition of a 'reasonable man' rests on this assumption: 'A reasonable man', he wrote, 'will seek to make his attitude exactly proportional to the balance of weight in the grounds of belief'.² Pepper assumed that the only real problem is the one an individual may face in getting his personal judgement into alignment with the 'balance of the weight' of the evidence; it seems never to have entered his head that, if thoughtful and intelligent people persistently differ in their assessments of the same evidence, the very notion of there being such a balance of weight may have to be abandoned, or at least radically revised.

But that thoughtful and intelligent people often do differ, and differ persistently, is obvious. Hence, instead of assuming uncritically that reasonableness is a viable concept, we ought rather ask ourselves two sobering questions: first, and as a matter of fact, why do reasonable people differ in their assessment of the same evidence; and second, and as a matter of principle, can such differences always be resolved?

Since these are very large questions it will probably be strategically sound to approach them via a concrete case. In this connection one thinks of Julian Jaynes' recent book³ on the bicameral mind; it is certainly one about which thoughtful and intelligent people differ. I propose therefore to examine the cognitive claims made in this book and to ask

- 1 The original version of this paper was prepared for a colloquium at Caltech in November 1978 in which Mr. Jaynes took part. In a considerably expanded form it was circulated as Humanities Working Paper No. 23, and was later presented to the Interpretive Studies Seminar at Caltech. I am most grateful to the members of the seminar, and more particularly to David Alexander, John Benton, Bruce Cain, Max Delbrück, Charles Young and Eran Zaidel, for helpful criticisms.
- 2 Stephen C. Pepper, *World Hypotheses*, Berkeley and Los Angeles 1961, p. 12. It would seem that reasonableness depends on achieving a ratio of unity between the degree of one's belief in any proposition *P* and the evidence for *P*. One is 'unduly credulous' if the ratio is much more than unity and 'excessively skeptical' if the ratio is much less.
- 3 J. Jaynes, *The Origin of Consciousness in the Breakdown of the Bicameral Mind*, Boston 1977.

why these claims are assessed so differently.⁴ Accordingly, this paper is divided into two parts. Part I sets out Jaynes' major contentions and the evidence he adduces to support them. Part II is a venture into the sociology of knowledge—or more precisely, the sociology of belief. It seeks to expose some of the extra-cognitive elements that were at work in the fashioning of Jaynes' argument, and it attributes the varied assessments of that argument to differential responses to these elements.

I

Jaynes' major contentions are, first, that four great changes—changes not previously noted by any historian of culture—occurred towards the end of the third millennium B.C. and second, that taken together, these changes account for the leading features of post-third-millennium culture.

One of these four changes was a radical transformation of human experience from a completely unthinking, stimulus-response type of behaviour to the self-conscious, reflective type of experience that we are familiar with today.

A second change was a change in the mode of social control—from an automatic control (never before identified by any student of social or political behaviour, but analogous to the control a hypnotist exerts over his patients) to the nonautomatic kinds of control (e.g., by threat, cajolery, bribery, or rational persuasion) that we know today.

The third change was a change in the incidence of what Jaynes calls 'florid, unmediated' schizophrenia: before the end of the third millennium B.C. 'everyone was schizophrenic' (p. 405), and, what is more, everyone was schizophrenic all the time (p. 140). Thereafter the incidence greatly declined.

The fourth change was a change in the function of that area on the right hemisphere that corresponds to Wernicke's area on the left hemisphere.

Since it appears—judging both by the title of the book and also by the amount of space devoted to 'the origin of consciousness'—that this is the centrepiece of Jaynes' case,⁵ I shall begin by examining his claims

⁴ There are many reasons other than a favourable assessment of a book's cognitive claims for welcoming it. One might, for instance, regard it as giving a needed jolt to complacent establishment types. Peter Brown, for instance, has recently praised *Hagarism* by Patricia Crone and Michael Cook in such terms: 'One can only hope that this daring thesis will be the object of prolonged scholarly debate; and that this debate, while it is likely to reject Crone and Cook's conclusions, will respect the problems to which they have drawn attention. For it is the problems raised in *Hagarism*—not the flimsy and precipitate solutions that the authors propose for these problems—which make this a book of genuine historical importance. *Hagarism* is a salutary reminder of how much still needs to be explained in the origins of medieval Islam' (*NYRB*, 26, 1979, 2, 32). My emphasis throughout is on differing estimates of the cognitive claims made by Jaynes, not on other possible reasons for approving or disapproving his book.

⁵ The primacy of the origin of consciousness in Jaynes' thinking was confirmed by him at the Caltech colloquium (see note 1).

Julian Jaynes and the Bicameral Mind 155

regarding the nature of this supposed transformation in the quality of human experience.

An introductory chapter reviews earlier answers to the question, 'what is consciousness?'—answers which Jaynes rejects as wholly inadequate. From this spectacular history of failure (I agree with Jaynes about the inadequacy), one might reasonably conclude that the question has been incorrectly posed, that it is incapable of being answered in the form in which it has been asked. But Jaynes does not consider this possibility; instead, the chapter ends with the observation, 'We must therefore try to make a new beginning by stating what consciousness is' (p. 18), and Chapter 2 is duly devoted to setting out the answer that, as he holds, really and finally explains not only what consciousness is but whence it comes.

Consciousness, according to Jaynes, is a 'creation' of language (p. 52), and specifically a creation of metaphor. Thus 'metaphor generates consciousness rather than describes it. . . . Consciousness is the work of lexical metaphor' (p. 56). What Jaynes means by 'description' (i.e., description by means of metaphor) is not, I think, in dispute. He remarks that 'most of the errors about consciousness. . . have been errors of attempted metaphor' (p. 53), and he specifically mentions the tendency to think of consciousness spatially, e.g., as a thing somehow located inside our heads. Who can disagree? But indictment of description by means of metaphor is hardly a novelty: Writers as diverse as Kierkegaard and Nietzsche, Brentano and Sartre, Husserl and Russell have all pointed out that it is misleading to spatialize consciousness and that our conclusions about the self are (as Nietzsche remarked) 'formulated out of our grammatical custom'.⁶

But the question is how Jaynes distinguishes 'generation by metaphor' from 'description by metaphor'. He acknowledges that what he says about generation by metaphor is 'fairly dense' (p. 54) and 'difficult and overtly diffuse' (p. 55). It is certainly all of that, but my complaint is rather that his account of the generation of consciousness by metaphor is itself metaphorical, and that when we manage to unpack the metaphor we find a spatialized consciousness not very different from the spatialized consciousness that he has earlier explicitly, and rightly, rejected.

What meaning does Jaynes attach to the expression 'metaphor generates consciousness'? More than once he points out that all of our language is deeply metaphorical, that even the most seemingly abstract terms (e.g., 'force', 'inertia', 'field', 'acceleration') have their roots in concrete metaphors: 'Abstract words are ancient coins whose concrete images in the busy give-and-take of talk have worn away with use' (p. 51), thus characteristically cloaking a truism in a glittering metaphor.

6 F. Nietzsche, *Beyond Good and Evil*, tr. M. Cowan, Chicago 1955, p. 19.

Speaking picturesquely, and very loosely, we might allow ourselves to say that metaphor created the language in which physicists now talk about nature; analogously, we might say that metaphor created the language in which psychologists talk about consciousness. So far, so good. But does Jaynes claim that metaphor created the nature that physicists talk about in their formerly metaphorical language? What he holds to be the relation between language and nature I cannot say. But as regards language and consciousness his position is plain, if puzzling: metaphor not only created the language in which we talk about consciousness, it also created the consciousness that is talked about in this language.

How can this seem plausible to Jaynes? The answer is that he believes language (and so metaphor) has a special relation to what he calls 'our immediate experience'. Certainly we can agree that the experiential field of people who possess language is very different from the experiential field of people (e.g., babies) who lack language: the experiential field of people without language is presumably much less highly articulated than the experiential field of people with language. But it is surely structured to some extent, even though these structures are rudimentary—babies, for instance, recognize their mothers and their bottles. So, too, most of us would say, regarding the consciousness of people without language: though it is much less highly articulated than the consciousness of people with language, it too contains structures, more or less paralleling the contrast between the contents of the experiential fields of people without, and people with, language.

Jaynes disagrees. Though he would presumably allow that the experiential field of people without language is more or less articulated into contents like 'parents' and contents like 'bottles', he holds that their field does not contain any distinction, in however rudimentary and wavering a form, between 'me' and 'not-me'; that is to say people without language are not conscious at all. Consciousness is not a matter of degree of complexity, as parents and bottles are matters of degree of complexity. There is a radical difference in kind between the pre-language experiential field and the post-language experiential field. Thus, at least according to Jaynes, language brought one absolutely new feature into the world—the distinction, within the field, between what is outer and what is inner, between what is experienced and the experiencing of it.

That is Jaynes' claim. Is there any evidence in the behaviour of babies and other people without language to support this claim? Jaynes cites none. Instead, he relies on a highly metaphorical account of the 'creative' power of metaphor, and when we unpack this metaphor about metaphor we are left with 'description by metaphor'—'generation by metaphor', as a special kind of linguistic activity, has disappeared.

Let us start the unpacking process by noting that it isn't language that creates a metaphor; it is people, using language, who create things, and

Julian Jaynes and the Bicameral Mind 157

among the things they create are verbal tokens which serve to mark various features of the experiential field. When it happens that somebody first notes a similarity between two features of the experiential field that are otherwise quite dissimilar, the token he uses to mark this similarity is called a metaphor. If somebody notes (an example of Jaynes') that his love is in some respects like a tin smith's scoop, he can do so only because there are two different things in his experiential field, his love and the scoop; if his love and the scoop are not, both of them, in his field, he cannot observe the similarity between them which he now marks with the metaphor (i.e., the verbal token). It is true that, given the sophisticated linguistic resources—semantical and grammatical—of somebody like Jaynes, this metaphor, by likening love to a scoop, articulates and brings into focus certain aspects of love that might otherwise escape notice. But that is all there is to the 'generation' of love 'by metaphor'; it is only 'description' of love 'by metaphor', itself dressed up, for a change, in metaphorical language.

This is also true for the supposed generation of consciousness by metaphor. When Jaynes tells us that language creates consciousness as an 'analog' of experience (an analogue in the same sense that a map is an analogue of the land mapped by the map), we have to remind him that language doesn't create anything, except metaphorically. Rather, people find analogues (read 'similarities') and mark these similarities by metaphors. Somebody—perhaps some early user of language in the remote past of the race—noted that consciousness is in some respects like physical space, just as Jaynes, much more recently, noted that love is in some respects like a scoop. And just as love and the scoop must both have been in Jaynes' experiential field for him to be able to note the analogue between love and the scoop, so consciousness and physical space must both have been in that early man's experiential field. And just as Jaynes explicated the details of the analogue between love and the scoop—the enduring careful shape and hidden shiningness and holdingness of a lasting love deep in the heavy manipulable shortness of mounding time' (pp. 57-58)—so he has also explicated the details of the analogue between consciousness and space:

Subjective conscious mind is an analog of what is called the real world. It is built up with a vocabulary or lexical field whose terms are all metaphors or analogs of behavior in the physical world. . . .

And the adjectives to describe physical behavior in real space are analogically taken over to describe mental behavior in mind-space when we speak of our minds as being 'quick', 'slow', 'agitated'

As with a real space, something can be at the 'back' of our minds. . . . In argument we try to . . . 'find a common ground', or 'point out', etc., all actions in real space taken over analogically into the space of mind. [P. 55.]

But it is Jaynes who is doing this job of explication, not metaphor. His—or, if one likes, early man's—metaphor only articulates an ob-

served similarity of two features of the experiential field; it did not generate one of them (consciousness) out of whole cloth.

So this is what the great distinction between 'description by metaphor' and 'generation by metaphor' comes down to: when ordinary chaps today think about consciousness in terms of a spatial metaphor (i.e., describe by metaphor), that is egregious error. But when early man did exactly the same thing, we are to call it generation by metaphor and hail it as an epoch-making moment in the history of mankind.

This might be thought to be the last thing that must be said about metaphor in connection with consciousness. But there is another feature of metaphor—one that Jaynes overlooks—that is highly relevant to Jaynes' discussion. Let us begin again with similarities, and the verbal tokens by which we mark them. Similarities are there, as it were, waiting for someone who is perceptive enough to note them and describe them (not create them!) by means of metaphors. But once a verbal token enters into social discourse, and comes to 'stand for' some experienced similarity, it may eventually be substituted for it: the experienced similarity that the token represents may then drop out of experience and be replaced by the token. That is literalization.

In a word, a metaphor marks the discovery of a genuine similarity; it may, on occasion, create a false one through overextension. For instance, we may see that the gas tank of our car and our bath tub are similar in that neither contains liquid, and we may mark this similarity by means of the token 'empty'. If use of this token then causes us to overlook the fact that the gas tank contains gas fumes, overextension of the metaphor may let us in for a nasty surprise.⁷

If we judge that some metaphor (any metaphor) has become irremediably overextended in this way—if we fear that we cannot escape the language trap of taking this metaphor literally—then we might decide to abandon the expression that is the token of the metaphor. Let us, we might urge, cease to talk about 'empty' gas tanks; that sort of talk is misleading. This is my recommendation regarding the term 'consciousness'. I think it should be dropped from any rigorous vocabulary; it has been ruined by our tendency to reify it, to take a spatial metaphor literally. The seductive power of this particular language trap is demonstrated, I think, by Jaynes' own discussion: though he *tells* us that consciousness is not a thing, only an 'operation', he slips back into thinking about it as if it were a thing—hence the pseudo-questions and pseudo-answers with which the chapter is concerned: What 'is' consciousness? (Answer: an analogue of physical space.) What is 'its' cause? (Answer: language.)

Well, then, if we are to abandon the term 'consciousness', what term, or terms, are we to use in its place? We want terms, if we can find them,

⁷ The example is Whorf's.

Julian Jaynes and the Bicameral Mind 159

that are less misleading because they are more resistant to literalization. I suggest that what Jaynes is really concerned with is what I would call 'deliberation', or 'reflection', or 'consideration of alternative possibilities'. We all know what such terms refer to, and would agree, I believe, that much of the time, people do not deliberate, reflect, or consider alternatives. Much of the time—most of the time perhaps—people respond, unreflectively, with learned routines, to perceived cues. These responses may be short behavioural outcomes or they may consist of long, complex serials; people even acquire meta-routines that can be brought into play (without reflection) when some serial breaks down.

Jaynes does not of course dissent from any of this, though he uses other terms to describe the process:

In driving a car, I am not sitting like a back-seat driver directing myself. . . . I am related to a world I immediately obey in the sense of driving on the road and not on the sidewalk. And I am not conscious of any of this. And certainly not logical about it. I am caught up, unconsciously enthralled, if you will, in a total interacting reciprocity of stimulation that may be constantly threatening, or comforting, appealing or repelling, responding to the changes in traffic and particular aspects of it with trepidation or confidence, trust or distrust, while my consciousness is still off on other topics. [Pp. 84-85.]

But now, suppose it happens that on some occasion none of the routines or meta-routines that are a part of our behavioural repertoire is appropriate. Then, characteristically, we begin to reflect; we consider alternatives. These alternatives are optional scenarios; they may be 'sketchy' or they may be worked out in detail; they may have very short, or very long time horizons. They may involve projections of ourselves into various possible futures which we compare. One can certainly, if one likes, characterize these 'operations' by such terms as 'narratization', an analogue 'I', and a metaphoric 'me' (pp. 59-65), though I don't think that much is gained by using them to replace such familiar terms as 'planning', 'calculating', and the like. In any case, however, we are now finally in a position to reformulate Jaynes' central contention without reference to all that talk about 'the generation of consciousness by metaphor'. His thesis is that before the end of the third millennium people never, in any circumstances whatsoever, reflected (however briefly); they never considered alternatives (however sketchy and with however short time horizons).

That is Jaynes' hypothesis. Before we consider the evidence which, in his view, validates it, we must briefly examine the other three postulated changes and ask how he conceives them to be related to this change in the structure of human experience which we have been discussing. Consider first the claim that social control was effected by means of hallucinated voices. Some such hypothesis seems essential if delibera-

tion (Jaynes' 'consciousness') originated at the time and in the way he claims. Deliberation (the projection of alternative scenarios) commonly occurs, even in our modern societies, only when some behavioural routine breaks down and no appropriate meta-routine is immediately available. Breakdowns of behavioural routine must have occurred in pre-third-millennial times. Why, then, did not deliberation ensue? Obviously Jaynes must find some reason and his explanation is that at the very moment a breakdown occurred and before deliberation had time to begin, a hallucinated voice spoke to pre-third-millennium man and told him exactly what to do.

The voice-hallucination hypothesis is also an answer to an otherwise insoluble puzzle about the nature of social control in pre-third-millennium times. All forms of social control that we know—such as threat, bribery, cajolery, and rational persuasion—presuppose some capacity, however slight, for reflection and deliberation. That is, to use Jaynesian language, they all presuppose some form, however fugitive, of an 'Analogic "I"' and a 'Metaphoric "Me"'. Thus his account of the origin of deliberation requires that, until he is ready to have deliberation enter the scene, there must be a form of social control that is as automatic and as unthinking as are routine responses to sensory cues. A voice believed to be divine and which one therefore obeys instantly and without question is exactly the kind of social control required in these circumstances, and it is difficult to think of anything else that meets the need.

Next, as regards the hypothesis that, before the end of the third millennium, 'everyone' was schizophrenic: though this is discussed as if it were a third great change that occurred simultaneously with the origin of deliberation ('consciousness') and the decline in the incidence of voice hallucination, a little consideration shows that it is simply a dramatic, and therefore rhetorically effective, restatement of the first and second hypotheses taken together. That this is the case follows from Jaynes' description of 'florid unmediated' schizophrenia: its symptoms, he writes, 'are primarily the presence of auditory hallucinations . . . and the deterioration of consciousness . . . namely, the loss of the analogue "I", the erosion of mind-space, and an inability to narratize' (p. 408). It may indeed be the case that the first and second hypotheses, if confirmed, throw some light on the nature of 'this most common and resistant of illnesses', for schizophrenia would then turn out to be, as Jaynes says, 'a vestige of bicamerality, a partial relapse to the bicameral mind' (p. 405). But the existence of the disease in historic times throws no light on the changes that are supposed to have occurred at the end of the third millennium. Rhetoric apart, then, the formulation of those changes as involving a radical decline in the incidence of schizophrenia adds nothing to the claims already stated in the first two hypotheses.

Julian Jaynes and the Bicameral Mind 161

Finally, there is the neurological hypothesis regarding the supposed transfer of dominance (in right-handed people) from the right to the left hemisphere. To ask whether certain functions could have been localized in the way Jaynes claims would be to raise a question in neurophysiology that I am not competent to deal with. It is possible however to pose another, and more basic question: what does Jaynes hold to be the relation between 'brain structure' on the one hand, and 'psychological phenomena', on the other? He seems to be of at least two, and possibly more, minds about this. At one point, he explicitly rejects what may be called an identity-theory of the relation between brain state and psychological state. There is 'a delusion', he writes, in 'the all-too-common and unspoken tendency to translate psychological phenomena into neuro-anatomy and chemistry' (p. 18). He also rejects, again quite explicitly, all forms of dualism: dualism, 'one of the great spurious quandaries of modern psychology', 'began its huge haunted career' with the pre-Socratics and included the 'arrogant assurances' of Descartes (p. 291). What are we left with? The only positive statement I can find is the following:

Of course it is extremely hazardous thinking to isomorphize between a conceptual analysis of a psychological phenomenon and its concomitant brain structure, yet this is what we cannot avoid doing. [P. 102.]

In any event, having assumed isomorphism despite its acknowledged hazards, Jaynes proceeds as if no hazards are involved. Thus he shifts back and forth all too easily from psychological language to physiological language, often in the same sentence. For instance:

In bicameral men... volition came as a voice that was in the nature of a neurological command. [P. 99.]

Each person had a part of his nervous system which was divine, by which he was ordered about... [Pp. 201-202.]

... an inaccessible voice in the nervous system... [P. 301.]

Is it possible that what corresponds to Wernicke's area on the right hemisphere 'looks down' on Wernicke's area on the left? [P. 349, note 12.]

... authority figures created by the nervous system out of the patient's admonitory experience... [P. 411.]

This is equivalent to writing: 'La donna ist very belle'—which is odd, but at least intelligible because Italian, German, English and French are roughly equivalent languages (to use Jaynesian terminology, we can 'isomorphize' from one to another). The sentences I have just quoted from Jaynes are equally odd, but there is no way of knowing whether or not they are intelligible, because it is far from settled that neurophysiology and psychology are equivalent languages. It is not known that the

explanatory paradigm used in physiology and formulated in its language is isomorphic with the explanatory paradigm used in psychology and formulated in its language.

Therefore, until their isomorphism is demonstrated, I think we must regard the origin-of-consciousness hypothesis and the voice-hallucination hypothesis as independent of the shift-of-dominance hypothesis. Evidence for the first two would not tend to support the latter, and evidence for the latter, if it could be found, would not support the former two. But how could evidence for a shift in dominance at the end of the third millennium be found independently of some assumptions regarding isomorphism?

I conclude that we are dealing at most with two, not four, 'great changes'. Let us now review the evidence Jaynes has marshalled to support his assertion that the changes described occurred at the end of the third millennium. This evidence can only be in the form of traces—documents, buildings, monuments, statues, figurines—that (1) have survived down to our own time, (2) can be dated back to times not later than those in which it is asserted that the great changes occurred, and (3) differ in specific features from any traces firmly dated later than those in which the changes are thought to have occurred. Only traces that meet these three criteria are available as evidence. For such traces actually to become evidence Jaynes must show that the specific features cited are 'better understood' on the origin-of-consciousness and the voice-hallucination hypotheses than on any alternative hypothesis.

I shall call this the positive case. It should be supplemented by what I shall call the negative case. This would consist in showing that the other features of these traces—and other traces not having these features—can be explained at least as well on the origin-of-consciousness and the voice-hallucination hypotheses as on any other hypothesis. The negative case is important because, unless it can be made, Jaynes will be left, at best, with two rival explanatory hypotheses, each of which covers only one part of the traces that are to be explained.

I shall take up first the negative case. What sorts of traces are, at least prima facie, inconsistent with Jaynes' hypotheses? One such class of cases is traces of large-scale organizations well before the end of the third millennium. It is possible, I suppose, to believe, without too much strain, that very small groups (e.g., nuclear families) could live on a day-to-day basis in a 'signal-bound [condition], that is, responding each minute to cues in a stimulus-response manner, and controlled by those cues' (p. 140). But with larger groups—even with groups no larger than an extended family—it takes extreme dedication to a hypothesis to continue to hold that the members could live out their lives in a completely signal-bound condition. And of course there are many traces from very early times of large organizations with fairly elaborate bureaucratic structures.

Julian Jaynes and the Bicameral Mind 163

Jaynes is of course not unaware of the intrinsic implausibility of the voice-hallucination hypothesis when it is applied to large-scale organizations. He suggests the possibility that in such societies the hallucinated voices may have been hierarchically organized. This sounds like an ad hoc hypothesis on a par with the epicycles that were introduced to support the geocentric hypothesis. And in any case, would hierarchically organized voices eliminate the possibility, even in very small societies, of the kind of breakdown in routines that, as we have already suggested, leads to deliberation? It seems highly probable that they would not. However well organized into hierarchies the voices might be, they could hardly fail, on occasion, to give conflicting instructions, or instructions that were inappropriate for the situation in which 'the gods' slaves' found themselves. It is a feature of instructions, however carefully designed, that they are too general to cover all possible circumstances: this is why judges and administrators are needed—to find the relevant rule and apply it to the particular case. Jaynes himself agrees that eventually organizations became too large for control by hallucinated voices, however hierarchically organized. He suggests the Intermediate Period in Egypt as the era when this occurred. I am simply pointing out that it must have occurred much earlier, if indeed there ever was a time when men were actually automata who 'responded every minute to cues in a stimulus-response manner'.

It is easy to think of other kinds of traces that are, at least *prima facie*, inconsistent with Jayne's hypotheses, but enough has already been said to show that the negative case is far from complete and that the *onus probandi* lies on Jaynes to complete it.

Next as regards the positive case: here we must distinguish between evidence for the origin-of-consciousness hypothesis and evidence for the voice-hallucination hypothesis. Generally speaking, the evidence for the former is documentary—for instance the text of the *Iliad* and 'the dispute of a man with his Ba'—while the evidence assembled for the latter consists in traces of buildings (any large structure that is not a dwelling and that has no other practical use), figurines (especially pop-eyed figurines), and burial practices (for instance, the burial of important dead as if they still lived).

I shall confine myself to discussing one example of the evidence assembled for each hypothesis. Consider then the cuneiform texts that contain signs which Jaynes believes are equivalent to such English expressions as 'speaking', 'uttering', 'hearing', and 'listening'. I submit that we do not know enough about the authors of these texts nor about the genres that the texts represent to be able to say whether the cuneiform signs are to be understood metaphorically or (as Jaynes would have it) literally. Our conclusion can only be 'not proven'.

Or consider the matter of burial practices. There are surely many reasons why survivors might have chosen to bury their dead as they did.

Jaynes himself suggests one such possibility—grief. Though he immediately rejects this possibility and concludes that survivors buried their dead as if they were alive because these survivors were literally hearing the voices of the dead who ordered their burial in this fashion, he admits that this conclusion is not ‘necessary’; it is only ‘consistent’ with the evidence (p. 165). But he soon forgets the tentativeness that (for once) he has affirmed and is back once again to categorical assertions: ‘has no clear explanation except . . .’ or its equivalent is the phrase with which he usually concludes his discussion of any trace that he regards as evidence for his hypotheses.

II

I have now given what I hope is a fair summary of Jaynes’ argument, and in doing so I have not tried to hide my own conclusions regarding his conclusions. My view is that the case is not only ‘not proven’; it is not remotely plausible. In a word, from my perspective, what Pepper would call ‘the weight of the evidence’ is badly out of balance. The ratio (as I see it) between the degree of Jaynes’ belief in his conclusions and the evidence he has presented for them is so much greater than unity that he seems to me ‘unduly credulous’. But I know that some readers find the book persuasive, or at the least suggestive. They may not perceive the weight as being exactly at unity (as Jaynes himself presumably does), but they don’t perceive it as being as far out of balance as others of us do. How can this be? In this section of the paper I shall try to account for these radically different perceptions of where the weight of the evidence lies.⁸

The account I shall give starts from an epistemological model that I have discussed in detail elsewhere,⁹ and applies this model to the case before us. I assume, then, that the objects of perception and cognition (what is perceived, what is understood) are the end results of processes

8 At one point during the Caltech colloquium (see note 1) Jaynes remarked that his conclusions regarding lateralization and localization of function are ‘astonishing’. ‘Incredible’, replied Eran Zaidel, the neurologist on the panel. That is the kind of difference in assessment that I seek to explain.

9 See W. T. Jones, *The Sciences and the Humanities*, Berkeley and Los Angeles 1965. Subsequent applications of the model to concrete cases of what I have called ‘nonterminating disagreements’ include ‘Philosophical Disagreements and World Views’, *Proceedings and Addresses of the American Philosophical Association*, 43, 1969-70, 24-42; ‘World Views: Their Nature and Their Function’, *Current Anthropology*, 13, 1972, 79-109; ‘Talking about Art and Primitive Society’, in A. Forge (ed.), *The Study of Primitive Art*, Oxford 1973, pp. 256-77; ‘World Views and Asian Medical Systems’, in C. Leslie (ed.), *Towards a Comparative Study of Asian Medical Systems*, Berkeley and Los Angeles 1965, pp. 383-404. More recently, and with the collaboration of William L. Faust, Margaret S. Faust and Molly Mason Jones, I have developed empirical procedures for testing hypotheses about the causes of nonterminating disagreements, but none of this material has yet been published.

Julian Jaynes and the Bicameral Mind 165

in which some 'foreground', i.e., some element in the experiential field, is mediated by a background structure consisting of memories, generalizations and attitudes. Thus the cognitive process involved in the assessment of the evidential value of some trace (say, the burial practice described above) may be represented as:

$$I = f(F, B)$$

where I = the interpretation, or assessment, of the evidential value of the trace

F = the trace, as foreground

and B = a background structure.

Obviously the background structures of people brought up within the framework of Western science (in contrast, say, to traditional Chinese science or to so-called 'primitive' ways of thinking about nature) will have many important features in common. Presumably all such people—including Jaynes himself, it is important to note—will have incorporated in their background structures a general belief to the effect that propositions that cannot be falsified and also propositions that have actually been falsified have no place in science. Further, and with regard to many particular propositions, such people probably agree which fall and which do not fall in these categories. Thus, over a wide range of experience one can expect agreement among scientifically trained people.

But—equally obviously—within such basic similarities of background structure differences of various kinds distinguish one individual from another and one group from another group. Among such differences the most notorious, perhaps, are racial, political and religious biases. Where such biases differentially characterize the individuals who happen to be assessing the same trace, differences in interpretation of that trace are likely to result. This may account for different interpretations of the role of heredity in IQ, the verification procedures provided in SALT II, and Cromwell's motives in voting for the execution of Charles I.¹⁰

What might be some of the differences in background structure that lead Jaynes and his critics to such radically different interpretations of the same evidence? Not such biases as those just mentioned. Rather, biases that are at once less readily detectable and more pervasive. For want of a better term to describe them I shall refer to them as differences in cosmological orientation, and I shall concentrate on three of these differences, without suggesting that these are the only such differences involved.

¹⁰ A report has survived—a trace—that Cromwell lifted the lid of the king's coffin and, gazing at his face, murmured, 'Cruel necessity'. This report was 'virtually accepted' by Gardiner and by Abbott; it has been characterized as 'incredible' by Maurice Ashley (*The Greatness of Oliver Cromwell*, New York 1957, p. 215).

First, then, Jaynes has a preference for abrupt, dramatic and radical change—a bias against gradualism and for discontinuity. This preference shows up in his references to ‘a huge alteration in human mentality’ and to ‘dramatic change’, in his liking for the expression ‘*de novo*’, and even in the title of his book. More importantly, I think this bias explains his persistence in using the expression ‘origin of consciousness’, despite his own observation that consciousness is not a ‘thing’, but an ‘operation’. Use of the term ‘consciousness’ comes naturally to Jaynes because it allows him to slip back into thinking of consciousness as a state of being and so makes it easy to think of it as coming into being all at once—an Athena springing full-grown from the brow of Zeus—whereas ‘deliberation’ (which I have argued is the term that best conveys what Jaynes is really talking about) has none of these connotations. Indeed, ‘deliberation’ naturally suggests continuities, not discontinuities, since it is a process that intrinsically admits of degrees—degrees of intensity, duration, and frequency of occurrence. Thus Jaynes’ talk about the origin of consciousness, instead of the origin of deliberation, not only reveals his bias, more to the point, it makes more plausible than it otherwise would be, his thesis that a major discontinuity in human culture occurred in the third millennium.

Discontinuity bias turns up in many other ways. For instance, I suggest that it accounts for Jaynes’ quite extraordinary hostility to Darwin. Since Darwinism reduces great changes to a very large number of very small changes occurring over immense times, it may be said to emphasize continuities instead of discontinuities; it substitutes a doctrine of gradualism for a doctrine of quantum leaps from one state to another wholly different one. No wonder, then, that Jaynes holds that Darwin ‘clouded the problem with his own naiveté’ (p. 9) and that the theory of natural selection is ‘a very suspicious totem of evolutionary mythology’ (p. 8).

Second, Jaynes has a bias against narration, spatialization, the ‘analogue ‘I’’ and the ‘metaphor ‘me’’—and a corresponding bias in favour of the unconscious state that he calls bicamerality—the state in which there is no reflection, no deliberation, and no interior debate because the gods speak directly to men. Thus, for instance, he writes with evident sympathy that the ‘yearning for divine volition and service is with us still’ (p. 313, see also p. 318); he contrasts the ‘sands of subjective uncertainties’ that characterize modern life (p. 320) with the ‘absolutes’ of life in bicameral times; and he describes the ‘nostalgic anguish’ that ‘subjectively conscious people’ feel for ‘the lost bicamerality’ (p. 297). And when he discusses modern schizophrenia it is in terms reminiscent of R. D. Laing’s thesis that schizophrenics are the only sane people in our insane modern world: the problem for the modern schizophrenic, Jaynes writes, is that his ‘relapse’ into bicamerality

Julian Jaynes and the Bicameral Mind 167

is only partial. The learnings that make up a subjective consciousness are powerful and never totally suppressed. And thus the terror and the fury; the agony and the despair. . . . The lack of cultural support and definition for the voices . . . produces a social environment that is a far different thing from the behavior of the *absolutely* social individual of bicameral societies. [P. 432.]

Jaynes, it would seem, holds that we would all be better off if 'everyone' were once again schizophrenic, if we could somehow return to a bicameral society which had not yet been infected by the disease of thinking too precisely on the event and the anguish of decision-making in an uncertain world.

The third, and last, cosmological orientation that I shall mention is Jaynes' desire for a sweeping, all-inclusive formula that explains everything that has happened, from the frequency of the occurrence of pop-eyed figurines and the oracle at Delphi to modern hypnosis and schizophrenia. Thanks, he believes, to his discovery of the changes that occurred toward the end of the third millennium B.C. it is possible to view world history as a 'drama', as an 'immense scenario' (p. 436). So, too, in a particularly revealing passage:

We are now at last in a position where we can look back and see the history of mankind of this planet in its proper values for the first time. . . . Our view of human history here must be that of a furthest grandeur. . . . We must try to see mankind against his entire evolutionary background, where his civilizations, including our own, are but as mountain peaks in a particular range against the sky. . . . [P. 317.]

Readers who share these three cosmological orientations are likely to assess Jaynes' argument very differently from those whose cosmological orientations differ from Jaynes'. Consider, for instance, Jaynes' hostility to Darwin, his rejection of natural selection, and his insistence that consciousness emerged suddenly from a type of human experience that was radically different in kind. Readers with a bias toward discontinuity (i.e., those whose background structures are similar in this respect to Jaynes') will tend to regard as plausible Jaynes' view that consciousness made an abrupt appearance in human history, even if they question the date which he assigns for this abrupt appearance. Readers with a contrasting bias toward continuity will tend to regard Jaynes' view as intrinsically implausible; they will favour the alternative hypothesis that no radical break occurred, either toward the end of the third millennium or elsewhere. Rather, they will hold that deliberation (whether measured in terms of frequency of occurrence, variety of alternatives examined, complexity of scenarios reviewed, and length of time-horizons taken into account) has been a component, however flickering, in human experience from the earliest times. Indeed, they will add, such a component can be inferred in the behaviour of the other primates and perhaps in other animals as well.

Or consider Jaynes' belief in an all-inclusive, complete explanation of the course of human history. Leaving aside for the moment the particular explanation he has proposed, many historians are profoundly skeptical of any single explanatory principle. Whereas writers like Toynbee and Sorokin and Spengler seem to share Jaynes' conviction that, though the world may have the superficial appearance of a mighty maze, it is nevertheless not without a plan, other historians see only 'the turbulent movement' of 'an unfathomed sea' and confess that, if there be a plan, it is 'on a scale beyond [our] comprehension'.¹¹ My thesis is that historians who share Jaynes' cosmological orientation will take more seriously any trace put forward as containing evidence for some 'immense scenario' than will historians who are skeptical of the very possibility of such scenarios. Given a difference in background structure (*B*) of this magnitude, traces (*F*) that are convincing evidence to some of Jaynes' 'immense scenario', will be far from convincing to others, and a nonterminating disagreement is likely to ensue.

But this contrast in cosmological orientation is of course by no means confined to historians. Scientists, too, are likely to have a differential tolerance for what all scientists recognize must be a certain looseness of fit between any given hypothesis and the evidence assembled for it. All may agree that, ideally, the evidence ought to fit so neatly that no alternative explanation is possible. All may also agree that this ideal is not attainable in real life. Again, all probably agree that, the more inclusive the hypothesis, the looser the fit is likely to be, but that, scope being a positive good, we ought to be willing to tolerate a certain looseness of fit in hypotheses of very great scope. But how much is scope worth? Differential evaluations of the worth of scope are responsible, in part, for the nonterminating disagreements that arise from time to time in all scientific disciplines.¹²

So much in general. If we now apply these observations to the examination that neurologists, archaeologists, linguists and psychologists might make of those parts of Jaynes' discussion that are relevant to their particular disciplines, I predict that they would make differential assessments of the evidential value of the traces he cites—assessments that reflect a differential tolerance for looseness of fit on the part of the scientists concerned. Nevertheless, and taking Jaynes' argument as a whole, I also predict that the reaction of most scientists would be

11 *Cambridge Medieval History*, vol. 8, p. 819; E. L. Woodward, *Short Journey*, London 1942, p. 141.

12 It was Claude Bernard, I think, who urged his fellow scientists to 'seek simplicity' but to 'mistrust' it. The fact that all scientists probably agree with this advice, and hold that they actually follow it, does not prevent disagreements from occurring. To the extent that scientists do a kind of cost-benefit analysis of the relative values of seeking and mistrusting simplicity they will compute the payoff differently if they weigh differently the probability of cosmological simplicity.

Julian Jaynes and the Bicameral Mind 169

skeptical if not hostile. If that were to prove to be the case, it would be due, I think, chiefly to the fact that by training, and even more perhaps by temperament, their mistrust of simplicity (following Claude Bernard's injunction) is far greater than Jaynes'.

So far I have been discussing the disagreements between Jaynes and his critics as if these were merely intra-scientific disagreements—disagreements among people all of whom accept the scientific world view but who reach different conclusions because of relatively minor differences in cosmological orientations. But this way of putting the matter hardly uncovers the full measure of the disagreement. After all, Galileo, Newton and Einstein certainly proposed sweeping generalizations, and scientific mistrust of simplicity has not deterred scientists from accepting them. It is not so much that Jaynes' generalizations are grand and overarching as that they are remote from any body of generalizations of lesser scope. Though they have the look of scientific generalizations, they are not scientific generalizations at all. They are products not of the scientific world view but of an altogether different world view, one that can best be characterized as poetic or religious in nature.

Jaynes' book is, of course, by no means the only work that looks to be, but in fact is not, an expression of the scientific world view. Consider, for instance, Laplace's claim that, for an intelligence that knew in one moment of time 'all the forces by which nature is animated . . . nothing would be uncertain and the future, as the past, would be present to its eyes'. It has recently been argued that:

this statement has the plausibility of a good passage of science fiction. Although we can show that it is absurd, the demonstration of its absurdity makes no impression on those scientists, or those nonscientists, who love its magnificent flourish, exalting a brilliant aspect of science.¹³

Whatever one thinks of this as a characterization of Laplace, it applies, I think, to Jaynes, though I would characterize his book as a prose poem rather than as science fiction. He is not so much offering us evidence as a religious or poetic vision; not evidence that we are to weigh, but a vision that may win us, not because it fits the facts more or less loosely, but because it satisfies some deep and extracognitive needs of our natures. But his book doesn't read like a prose poem; the rhetoric is quite different. We have, therefore, to account for all that neurophysiological, archaeological, and psychological talk; for the fact that evidence is seemingly presented and that conclusions are seemingly drawn.

This rhetoric is in fact an important element in the attraction the vision has for some readers. We can understand that attraction better if we see

13 Michael Polanyi and Harry Prosch, *Meaning*, Chicago 1975, p. 105.

that Jaynes' book belongs to the same genre as *A Study of History* and *The Decline of the West*, the genre that I call secular theology—'theology' because, like the great theological works of past ages, it presents a vision of the world as a whole. I call it 'secular' both because this vision is confined to what St. Augustine called 'the earthly city' and also because it is composed in a language that looks scientific, rather than in the language of ecstatic mysticism or of formal theology.

This certainly does not dissolve the disagreement between Jaynes and his critics but it does refocus it. The disagreement turns out to be not an intrascientific disagreement over whether the evidence Jaynes presents warrants the conclusions he reaches, but an extrascientific disagreement over how to evaluate a kind of writing in which what looks like evidence and conclusions is but the articulation of 'an immense scenario' which appeals, if it does appeal, because of its immensity and not because of the evidence for it. It is a disagreement, in a word, over how to evaluate secular theology.

My description of Jaynes' book as secular theology has already brought us to the final topic of this paper—the reasons for its success, despite its lack of scientific rigour. We now see that it is not despite, but precisely because of its lack of rigour that the book convinces those whom it convinces. We have seen that the book might be recommended for reasons having nothing to do with the cognitive claims it makes.¹⁴ But those who regard those cognitive claims as warranted are probably people who crave certainty, absolutes, final solutions, a meaningful world rather than a merely turbulent sea. On the other hand, they are people who are too much persuaded of the preeminence of science as *the* way to truth to accept any assertions that seem to lack the cachet of science. Secular theology is the literary form that reflects these two currents in our secularized modern culture.¹⁵

¹⁴ See note 4, above.

¹⁵ Most secular theologians would indignantly reject this account of how their writings function in generating belief; I am not at all sure that Jaynes would. My reason is a remark inserted at the very end of his book. This remark follows his account of what he calls 'scientism', a kind of thinking that is exemplified by such schools of thought as Marxism, psychoanalysis and behaviourism. All scientisms, according to Jaynes, have certain features in common: a 'rational splendor' that 'explains everything', a 'charismatic leader', a series of canonical texts, and the requirement of total commitment. These features, it is evident, all characterize Jaynes' own book. In fact, the writings I have called secular theologies are indistinguishable from Jaynes' scientisms, which, as he says, 'are clusters of scientific ideas which come together . . . into creeds or beliefs, scientific mythologies which fill the felt void left by the divorce of science and religion in our time' (p. 441). But now, having given what one would think is a devastating indictment of scientism, or secular theology, Jaynes casually remarks, 'and this essay is no exception'.

When during the Caltech colloquium I quoted this sentence to Jaynes and asked what we are to make of it, he evaded me. That he did so is hardly surprising, for he cannot want to face the implications of admitting that his book is a scientism. In the whole

Julian Jaynes and the Bicameral Mind 171

If I am right, then, one source of the book's appeal is that it is a new gospel, a world-picture startlingly different from any we are accustomed to and one in which everything has its secure place and all is accounted for. But that is not the book's sole source of appeal. It attracts, as well, those who resonate with the romantic primitivism expressed in Jaynes' longing for 'lost bicamerality'. Dislike or distrust of consciousness is almost a hallmark of modern culture. It began at least as long ago as Dostoevsky,¹⁶ it is strongly expressed in Nietzsche,¹⁷ it is a principal motif in Lawrence,¹⁸ it dominates Sartre—think of his anguish (not too strong a word) at his inability to collapse his *pour-soi* mode of being into the *en-soi*.¹⁹ And what is true of the high culture is as characteristic of low and middle-brow culture—Zen, transcendental meditation, sensitivity training and the other countercultural phenomena that arose in the nineteen-sixties and are still with us. People who participate in this climate of opinion are likely to resonate with the thesis that consciousness is a late, dateable, and on the whole regrettable arrival on the human scene.

Resonances of this kind, then, are important elements in the formation of belief. Belief is certainly on occasion generated by evidence, but it is seldom generated purely, or exclusively, by evidence. Given the looseness of fit that characterizes the relation between generalizations and data, something more than, or different from, evidence often tips the scales in the direction of one alternative rather than another. Usually the role of these extra-cognitive elements in inducing belief is not very noticeable. Jaynes' book is useful precisely because the role of such extra-cognitive elements as the three cosmological orientations I have been discussing is plain to see. This is why his book is a good case study in the sociology of belief.

book, up to this final page, he had written in a way that leads every reader to assume that Jaynes expects his assertions to be tested by scientific, not by visionary, criteria. If, as he was finishing it, it occurred to him that he had produced not a scientific treatise, but a visionary tract, no wonder that this discovery was acknowledged in such an off-hand way.

16 The protagonist of *Notes from the Underground* held that 'any sort of consciousness is a disease'.

17 For instance, in 'On Truth and Life': 'In some remote corner of the universe, poured out and glittering in innumerable solar systems, there once was a star on which clever animals invented knowledge. That was the haughtiest and most mendacious minute of "world history"—yet only a minute'.

18 Thus Birkin in *Women in Love*: 'You yourself, don't you find it a beautiful clean thought, a world empty of people, just uninterrupted grass, and a hare sitting up?'

19 In *Nausea*, Roquetin says, 'I, too, wanted to *be*', and in *The Reprieve*, Daniel, a homosexual, exclaims, 'Just to be in the dark, at random! To be homosexual just as the oak is an oak. To extinguish myself'.