Experiential facts?

plans, stores them in memory, and keeps track of where the execution of each one stands.

I suggest that the escape from stimulus-response behavior makes consciousness necessary, a keeping track in memory of internal rather than external controls on behavior. It is this continuous plan-monitoring function, and nothing more or less, that we define as consciousness. Thus consciousness is emergent from the process of driving behavior from internally held plans. Consciousness is not an object in itself, but a side effect of other neurological operations. In defining consciousness in this way we also redefine psychology, from a study of stimulus-response contingencies to a study of the plans that drive behavior.

There is no room for a "Cartesian Theater" in this conception, but the function of consciousness clearly fits with a parallel "multiple-drafts" notion, for many plans exist simultaneously (everything from short-term plans such as cooking dinner to long-term ones such as earning a Ph.D.). The control of behavior is a giant juggling act. Existing plans are evaluated along with incoming sensory information and the execution of one plan wins out. At the same time, new plans and new subparts of existing plans are created.

This interpretation of the function of a conscious mind has consequences in several illusions where the differences between consciousness and reality are particularly evident. One such consequence is the illusion of the knife-edge of time discussed by D & K. There is no need to micromanage temporal relationships in a range where time distortions in the incoming sensory channels and the motor apparatus begin to become significant. But the concept of the knife-edge precision of definition is necessary to assign a temporal order, however arbitrary, to events and actions. As D & K point out, the temporal tag is not itself a temporal event; it is useful in the planning and plan-executing processes, not in the real-time operation of the brain.

A second illusion, not discussed by D & K but illuminating in this context, is the feeling that the visual field presents a detailed and veridical representation of the surrounding world. Everyone shares an almost irresistible introspection that the visual world present in consciousness is, for example, in full color and sharp focus. Yet we know that reasonable color coding exists only in the central 30° or so of the retinal image, and that sharpest focus and high-acuity imaging occupy only a tiny region in the fovea. We see not the retinal image, but some idealized combination of sensory information, memory, and assumption combined so seamlessly that we are unaware that most of what we perceive isn't actually available in the retinal signal. It is just this composite that is useful in making decisions based on visual information – the immediately present visual image is just a processing stage, a small part of the available information.

A third illusion is the distal reference that characterizes both sensory and motor operations. We perceive objects in the world, not in the eye or ear, and we feel objects, not deformations on the skin. Awareness arises only where it is functionally advantageous, at a level of coding where sensory and motor processes are coded as common, distally oriented events (Prinz 1991). A wealth of empirical work (summarized by Bridgeman 1990) is now available that distinguishes what processing is available in consciousness and what is not.

It may seem ironic that the processes of creating plans, accessing memory, and keeping track of everything are themselves unconscious, that we are unaware in consciousness of the functions that support consciousness. But if we think of consciousness as a result of planning capabilities, not as a system in itself that must be modeled, there is no reason why consciousness should appear at a mechanism level and there is no mechanism to make that possible. Consciousness is not a monitor of mental life but a result of mental operations separated from the immediate sensorimotor world.

Experiential facts?

Andy Clark

School of Cognitive and Computing Sciences, University of Sussex, Falmer, Brighton BN1 9QH, England

Electronic mail: *andycl@cogs.sussex.ac.uk*

In this timely piece, Dennett & Kinsbourne (D & K) attack what they term the image of the Cartesian Theater. The key move in their argument is to insist that just as judgments about, for example, redness do not require corresponding brain events that are actually red, so, too, judgments about temporal sequence need not involve the construction, somewhere in the brain, of an actual sequence of brain events that are temporally related in just the same way as the events which figure in the judgment (see e.g., sect. 1.1, para. 6; sect. 2.1, para. 6). To suppose that judgments of sequence must depend on the absolute sequence of brain events at some (functional) point where "it all comes together" is a mistake that, it is claimed, can make a variety of phenomena involving subjective judgments of temporality seem needlessly anomalous. I believe D & K are right to reject this strong image of the Cartesian Theater, but their argument goes further, for they then go on to deny the distinction between:

a. cases where an agent actually has a subjective experience of some event (like seeing a woman without glasses) but later comes to judge that the woman wore glasses all along (the "Orwellian" story – sect. 2.2., para. 7), and

b. cases where the agent's original experience was of seeing a woman with glasses (even though the woman in question was not wearing any), and this experience is accurately recalled in the later judgment that she had glasses on (the "Stalinesque" story, sect. 2.2, para. 8).

It is not obvious (to me at least) why this distinction needs to be denied. For such a denial is not *forced* upon us when we give up the image of the Cartesian Theater. We may agree that later judgments of the temporal order of events need not be grounded in a kind of action replay in which we construct brain events of matching temporality. But we may *still* believe that there were facts about the immediate contents of conscious experience so that it can (for example) simply be true (or false) that at time *t*, you had the experience of seeing a woman without glasses. Such conscious states may surely form an absolute temporal sequence independent of the agent's propensity to later judge that given conscious states did or did not occur. And this is, on the face of it, all that is needed to justify the Orwellian/Stalinesque distinction. In short, I don't see why recognising the errors D & K point out undermines the idea of an absolute timing of conscious contents or of absolute facts about conscious contents. Such facts could be quite independent of our later judgments, and be facts nonetheless.

D & K's position is somewhat clarified by the example of metacontrast. Here, faced with the question "Did conscious perception of a disc occur?" they decline to answer, saying that "information about the disc was briefly in a functional position to contribute to a later report, but this state lapsed" (sect. 2.2, para. 29). The question of whether the disc perception was ever conscious is one that D & K claim is opaque to both the agent and to any outside observers. Probe the agent at different times and you will get different answers. Both Stalinesque and Orwellian stories are, it seems, "consistent with whatever the subject says, or thinks, or remembers" (sect. 2.2, para. 36). But this is surely only true if "thinks" here means "later judges to be the case." If we believe that there are facts about conscious contents and that such facts are in principle independent of later verbal reports, the distinction can be maintained. Perhaps it is the idea of facts about conscious mental experiences independent of facts about later verbal judgments that D & K really seek to displace?

I suspect that this is indeed the case and that D & K really

want to cast doubt on the very idea of an experiential fact – a fact about the content of conscious experience at a given moment. The arguments concerning temporal and spatial smear (sect. 1.1) suggest that they wish to reject the very idea of a single conscious observer as a locus of experiential facts. But nothing in the explicit argument seems to justify this radical conclusion. We could grant that a variety of brain states (spatially distributed) could be implicated in the construction of immediate conscious contents and yet still discover that there is some functional property (e.g., of synchrony of neural activity in certain regions), which is both necessary for a content to become consciously known and yields an absolute temporal order of experiences (with specific contents) – an order that need not, however, be preserved in later judgments about the order. I cannot see that this possibility is ruled out by anything that Dennett and Kinsbourne tell us.

To sum up, the move from the (proper and important) rejection of the idea that judgments of temporality require a matching temporal sequence of brain events to the denial of the Orwellian/Stalinesque distinction looks problematic. The transition could be oiled by some radical views about the nonexistence of experiential facts or the relation between conscious content and verbal report. If there is indeed such a hidden agenda, it should come on stage for the curtain calls.

The selfless consciousness

Antonio R. Damasio

Department of Neurology, University of Iowa College of Medicine, Iowa City, IA 52240

Electronic mail: *cmdardpg@uiamvs.bitnet*

I enjoyed reading Dennett & Kinsbourne's (D & K's) target article. It provides interesting ammunition against the intuition that consciousness depends on a single brain locus where multifarious information comes together, in spatial and temporal terms. This notion, which D & K modestly refer to as a "prevailing view," is far worse than that: It informs virtually all research on mind and brain, explicitly or implicitly, and is certainly the common sense concept of the nonscientist and nonphilosopher in the street. My other comments are as follows:

(1) The evidence presented by D & K draws on cognitive science and is damaging enough to the Cartesian Theater model. I think, however, that D & K could have made an even stronger case by using evidence from experimental neuroanatomy, neurophysiology, neuropsychology, and computational neuroscience. Elsewhere I have analyzed part of that evidence to construct an argument against one or even a few "integrative" brain sites (Damasio 1989a; 1989b; 1990). For instance, there is no neuroanatomical structure in the cerebral cortex to which signals from all the sensory modalities that may be represented in our experience can converge, spatially and temporally. The entorhinal cortex and the hippocampus might be candidates for that sort of "integrative" role but they do not pass the necessary anatomical tests. Also, we know for certain that they cannot do the job because patients in whom such structures are destroyed bilaterally (e.g., patients Boswell [see Damasio et al. 1989] and H. M. [see Corkin 1984]) do not have a disturbance of consciousness in the sense discussed by D & K. (It can be argued that Boswell's highest level of self-consciousness is not intact since he cannot access a large body of unique memories from his past, but it is clear that he deals quite self-consciously and appropriately with the universe, at categorical level). The prefrontal cortex, another region associated with consciousness in the minds of most people that have ever thought about the brain, is an even less adequate candidate than the entorhinal cortex for the "integrative" locus underlying a Cartesian Theater. It provides many anchor points for signals hailing from various sensory streams and from the motor system, but there is no single site to which "representations" can cohere spatially and temporally. Extensive bilateral ablation of prefrontal cortices in humans does not preclude basic consciousness, although, again, we have argued that the highest levels of self-consciousness are not possible without these structures.

(2) I had two problems with D & K's proposal. The first and most important is that the rejection of *one* biologically impossible Cartesian Theater does not amount to rejecting the sense of *one* self doing the experiencing. There are, without a doubt, neural systems whose operation generates the sense of self, and on the basis of which we construct the false intuition that *there is one brain site where experience happens.* A satisfactory model of consciousness should indicate how the dis-integrated fragments operate to produce the integrated self. My impression is that the Multiple Drafts model is part of an alternative to the Cartesian Theater model but not a complete one. I would suggest that there are two necessary functions missing from the Multiple Drafts model without which I cannot fathom how consciousness, illusory phenomenon that it may be, will emerge. The first function is the sustained updating of critical sets of knowledge of the individual doing the experiencing. The sets encompass both taxonomically categorical levels of knowledge ("supraordinate" and "basic object"), as well as unique level ("subordinate" and autobiographical). The updated knowledge refers not only to the past but also to the future, that is, to memories of intended actions and plans. The second function is the sustained monitoring of somatic states of the experiencer, to include both visceral and musculoskeletal sectors of the organism. I suspect that the updating of previously acquired knowledge is implied in the Multiple Drafts hypothesis, but I saw no reference to the possible role of somatic states. I do not believe consciousness is possible without having something like the multiple drafts of D & K *referred to the somatic base of the experiencer.* Only awareness, in the sense used by Crick and Koch (1990), might be possible without a somatic reference.

I am persuaded that the multiple drafts mechanism alone will produce a selfless, disembodied consciousness. Incidentally, selfless consciousness can *almost* happen in some circumstances. An example is anosognosia, a neurological condition caused by extensive parietal and frontal damage in the human right hemisphere. The patients are unable to monitor their somatic states comprehensively and become unconcerned with their medical problems and with their future implications. They can give evidence that some externally generated representations of their own body do not pertain to themselves. I usually teach about this condition by stating that the lesion has "chipped part of consciousness away," that many percepts and thoughts of these patients are no longer referred to their bodies (for a similar perspective on the neurobiological basis of the self, see Merleau Ponty or, more recently, Edelman, 1989). Deep level meditation is another circumstance in which consciousness loses itself, so to speak, and eventually dissolves (this can be achieved in certain forms of Buddhist meditation).

(3) My second problem has to do with the degree of dissolution of the Cartesian Theater in D & K's alternative. How dis-integrated, neurally speaking, need dis-integration be? I agree that there is no single Cartesian Theater, but I suspect that there may be *many* such theaters, or, to use my own metaphor, many stages on which relatively coherent drafts of ongoing neural activity play out, at slightly different times. The point here is that it is plausible that some components of our experience actually depend on a local integration of neural sets of activity. For example, under certain circumstances this might happen in primary sensory cortices as a result of synchronization generated by feedback.

My closing comment is about the connection between consciousness and the timing of neural events. This is an important issue and is finally receiving the attention it deserves. That time can provide the illusion of a single place has been proposed by