

9

*The Autobiographical Self**Memory Made Conscious*

Autobiographies are made of personal memories, the sum total of our life experiences, including the experiences of the plans we have made for the future, specific or vague. Autobiographical selves are autobiographies made conscious. They draw on the entire compass of our memorized history, recent as well as remote. The social experiences of which we were a part, or wish we were, are included in that history, and so are memories that describe the most refined among our emotional experiences, namely, those that might qualify as spiritual.

While the core self pulses away relentlessly, always "online," from hint half-hinted to blatant presence, the autobiographical self leads a double life. On the one hand, it can be overt, making up the conscious mind at its grandest and most human; on the other, it can lie dormant, its myriad components waiting their turn to become active. That other life of the autobiographical self takes place offscreen, away from accessible consciousness, and that is possibly where and when the self matures, thanks to the gradual sedimentation and reworking of one's memory. As lived experiences are reconstructed and replayed, whether in con-

scious reflection or in nonconscious processing, their substance is reassessed and inevitably rearranged, modified minimally or very much in terms of their factual composition and emotional accompaniment. Entities and events acquire new emotional weights during this process. Some frames of the recollection are dropped on the mind's cutting-room floor, others are restored and enhanced, and others still are so deftly combined either by our wants or by the vagaries of chance that they create new scenes that were never shot. That is how, as years pass, our own history is subtly rewritten. That is why facts can acquire a new significance and why the music of memory plays differently today than it did last year.

Neurologically speaking, this building and rebuilding job occurs largely in nonconscious processing, and for all we know, it may even occur in dreams, although it can emerge in consciousness on occasion. It makes use of the convergence-divergence architecture to turn the encrypted knowledge contained in dispositional space into explicit, decrypted displays in the image space.

Fortunately, given the abundance of records of one's lived past and anticipated future, we do not need to recall all of them or even most of them, whenever our selves operate in autobiographical mode. Not even Proust would have needed to draw on all of his richly detailed and long-ago past to construct a moment of full-fledged self-Proustiness. Thankfully, we rely on key episodes, a collection of them actually, and, depending on the needs of the moment, we simply recall a certain number of them and bring them to bear on the new episode. In certain situations, the number of summoned episodes can be very high, a true flood of memories suffused with the emotions and feelings that first went with them. (One can always count on Bach to bring about such a situation.) But even when the number of episodes is limited, the complexity of memoranda involved in structuring the self is, to put it modestly, very large. Therein lies the problem of constructing the autobiographical self.

Constructing the Autobiographical Self

I suspect that the brain's strategy for constructing the autobiographical self is as follows. First, substantial sets of defining biographical memories must be grouped together so that each can be readily treated as an individual object. Each such object is allowed to modify the protoself and produce its pulse of core self, with the respective feelings of knowing and consequent object saliency in tow. Second, because the objects in our biographies are so numerous, the brain needs devices capable of coordinating the evocation of memories, delivering them to the protoself for the requisite interaction, and holding the results of the interaction in a coherent pattern connected to the causative objects. This is not a trivial problem. In effect, complex levels of autobiographical self—those that, for example, include substantial social aspects—encompass so many biographical objects that they require numerous core self pulses. As a consequence, constructing the autobiographical self demands a neural apparatus capable of obtaining multiple core self pulses, within a brief time window, for a substantial number of components and holding the results together transiently, to boot.

From a neural standpoint the coordinating process is especially complicated by the fact that the images that constitute an autobiography are largely implemented in the image workspaces of the cerebral cortex, based on recall from dispositional cortices, and yet, in order to be made conscious, those same images need to interact with the protoself machinery, which, as we have seen, is largely located at brain-stem level. Constructing an autobiographical self calls for very elaborate coordinating mechanisms, something that the construction of the core self can, by and large, dispense with.

By way of a working hypothesis, then, we can say that constructing the autobiographical self depends on two conjoined mechanisms. The first is subsidiary to the core self mechanism and guarantees that each biographical set of memories is treated as an object and made conscious

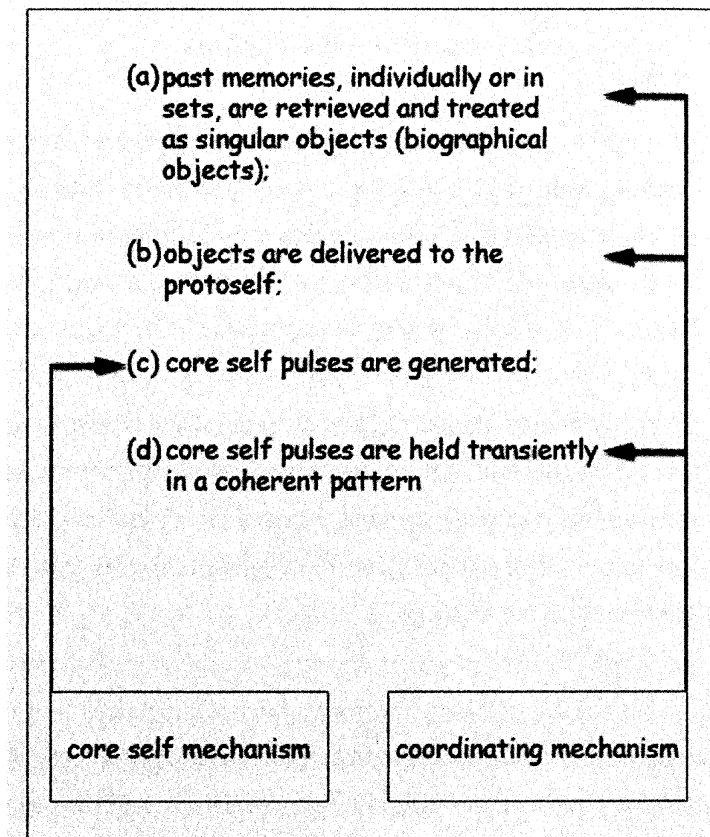


Figure 9.1: The autobiographical self: neural mechanisms.

in a core self pulse. The second accomplishes a brain-wide operation of coordination that includes the following steps: (1) certain contents are evoked from memory and displayed as images; (2) the images are allowed to interact in an orderly manner with another system elsewhere in the brain, namely, the protoself; and (3) the results of the interaction are held coherently during a certain window of time.

The structures involved in constructing the autobiographical self include all those required for the core self, in the brain stem, thalamus, and cerebral cortex, and, in addition, the structures involved in the coordination mechanisms discussed below.

The Issue of Coordination

Before I say one more word about coordination, I would like to make certain that my idea is not misinterpreted. The coordinating devices that I am postulating are *not* Cartesian theaters. (There is no play being performed inside them.) They are *not* consciousness centers. (There is no such thing.) They are *not* interpreter homunculi. (They know nothing, they do not interpret anything.) They are precisely what I am hypothesizing them to be and no more. They are spontaneous *organizers* of a process. The results of the entire operation *materialize not within the coordinating devices* but rather *elsewhere*, specifically, within the image-making, mind-generating structures of the brain located in both the cerebral cortex and the brain stem.

The coordination is driven not by some mysterious agent external to the brain but rather by natural factors such as the order of introduction of imaged contents in the mind process and the value accorded to those contents. How is the valuation achieved? Consider that any image being processed by the brain is automatically appraised and marked with a value in a process based on the brain's original dispositions (its biological value system), as well as on the dispositions acquired over lifelong learning. The marking stamp is added during the original perception and is recorded along with the image, but it is also revived during every instance of recall. In brief, confronted with certain sequences of events and a wealth of past knowledge filtered and marked by value, the brain's coordinating devices assist with the organization of the current contents. Moreover, the coordinating devices deliver the images to the protoself system and finally hold the results of the interaction (pulses of core self) in a transient coherent pattern.

The Coordinators

In the working hypothesis presented here, the first stage of the implementation of the neural autobiographical self requires structures and mechanisms already discussed for the core self. But there is something distinctive about the structures and mechanisms needed to implement the second stage of the process, namely, the brain-wide coordination described earlier.

What are the candidates for this large-scale system-coordination role? Several possible structures come to mind, but only a few can be seriously considered. An important candidate is the thalamus, a perpetual presence in any discussion of the neural basis of consciousness, specifically its collection of associative nuclei. The intermediate position of the thalamic nuclei, between the cerebral cortex and the brain stem, is ideal for signal brokering and coordination. Although the associative thalamus is busy enough constructing the background fabric of any image, it plays a very important, albeit perhaps not the lead, role when it comes to coordinating the contents that define the autobiographical self. I will say more about the thalamus and coordination in the next chapter.

What are the other likely candidates? A strong contender is a composite collection of regions in both cerebral hemispheres that is distinguished by its connectional architecture. Each region is a macroscopic node located at a major crossroads of convergent and divergent signaling. I described them as convergence-divergence regions or CDRegions in Chapter 6 and indicated that they are made of numerous convergence-divergence zones. CDRegions are strategically located within high-order association cortices but not within the image-making sensory cortices. They surface in sites such as the temporoparietal junction, the lateral and medial temporal cortices, the lateral parietal cortices, the lat-

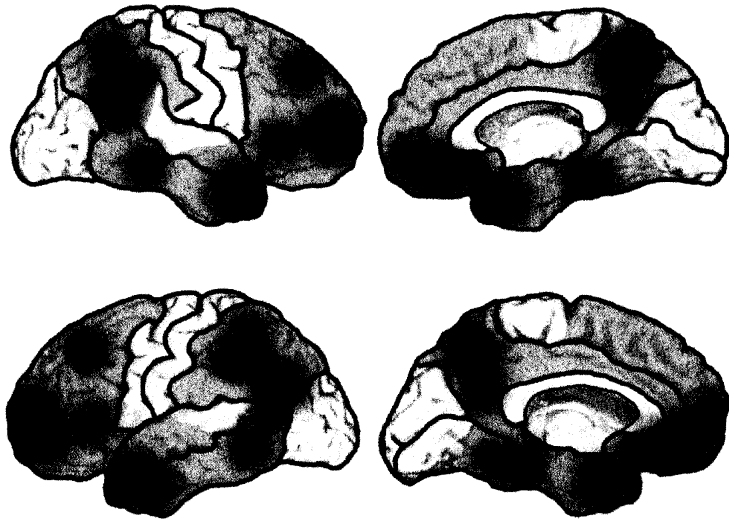


Figure 9.2: The task of coordinating the varied images generated by ongoing perception and recall is assisted by convergence-divergence regions (CDRegions), which are located within the nonmapped association cortices. The approximate location of the main CDRegions is suggested in the diagram (darkly shaded areas): the polar and medial temporal cortices, the medial prefrontal cortices, the temporoparietal junctions, and the posteromedial cortices (PMCs). In all likelihood, there are other such regions. Most of the CDRegions depicted in the figure are also part of Raichle's "default network" discussed later in this chapter. See Chapter 6 and Figures 6.1 and 6.2 for the architecture of these regions. See Figure 9.4 for connectional details of one CDRegion, the PMCs.

eral and medial frontal cortices, and the posteromedial cortices. These CDRegions hold records of previously acquired knowledge regarding the most diverse themes. The activation of any of these regions promotes the reconstruction, by means of divergence and retroactivation into image-making areas, of varied aspects of past knowledge, including those that pertain to one's biography, as well as those that describe genetic, nonpersonal knowledge.

Conceivably, the main CDRegions could be further integrated by long-range cortico-cortical connections of the kind first identified by

Jules Déjérine a century ago. Such connections would introduce yet another level of interareal coordination.

One of the main CDRegions, the posteromedial cortices (PMCs), appears to have a higher functional hierarchy relative to the others and exhibits several anatomical and functional traits that distinguish it from the rest. A decade ago I suggested that the PMC region was linked to the self process, albeit not in the role I now envision. Evidence obtained in recent years suggests that the PMC region is indeed involved in consciousness, quite specifically in self-related processes, and has provided previously unavailable information regarding the neuroanatomy and physiology of the region. (The evidence is discussed in the last sections of this chapter.)

The final candidate is a dark horse, a mysterious structure known as the claustrum, which is closely related to the CDRegions. The claustrum, which is located between the insular cortex and the basal ganglia of each hemisphere, has cortical connections that might potentially play a coordinating role. Francis Crick was convinced that the claustrum was a sort of director of sensory operations charged with binding disparate components of a multisensory percept. The evidence from experimental neuroanatomy does reveal connections to varied sensory regions, thus making the coordinating role quite plausible. Intriguingly, it has a robust projection to the important CDRegion that I mentioned earlier, the PMC. The discovery of this strong link occurred only after Crick's death and was thus not included in the posthumously published article that he wrote with Christof Koch, in which he made his case.¹ The problem with the claustrum's candidacy as coordinator resides in its small scale when we consider the job that needs to be performed. On the other hand, given that we should not expect any of the structures discussed earlier to perform the coordinating job single-handedly, there is no reason why the claustrum should not make a relevant contribution to the construction of the autobiographical self.

A Possible Role for the Posteromedial Cortices

We need additional research to determine the specific role the PMCs play in the construction of consciousness. Later in this chapter, I review evidence from varied sources: anesthesia research, sleep research, research on neurological conditions (ranging from coma and vegetative state to Alzheimer's disease), and functional neuroimaging studies of self-related processes. But first let's look at the PMC evidence that appears most solid and interpretable—evidence from experimental neuroanatomy. I'll speculate on the possible workings of the PMCs and on the reasons why they should be investigated.

When I proposed that the PMCs would play a role in generating subjectivity, there were two strands of thinking behind the idea. One strand concerned the behavior and presumed mental status of neurological patients with focal damage to this region, which includes the damage caused by late-stage Alzheimer's disease, as well as extremely rare strokes and brain metastases from cancer. The other strand related to a theoretical search for a brain region physiologically suitable to bring together information about both the organism *and* the objects and events with which the organism interacts. The PMC region was one of my candidates, given that it appeared to be located at an intersection of pathways associated with information from the visceral interior (interoceptive), from the musculoskeletal system (proprioceptive and kinesthetic), and from the outside world (exteroceptive). The factual strands are not in question, but I no longer see a need for the functional role I had envisioned. Still, the hypothesis prompted investigations that yielded important new information.

Making headway with the hypothesis was not easy; the main problem was that the neuroanatomical information available on this region was quite limited. Some valuable studies had begun to chart the connectivity of parts of the PMC,² but the overall wiring diagram of the region

had not been investigated. In fact, the region was known not by an umbrella term but rather by its component parts, namely, the posterior cingulate cortex, the retrosplenial cortex, and the precuneus. The PMCs, by whatever name, were definitely not yet on the radar of notable brain areas.

In order to explore the hypothesis that the PMC was involved in consciousness, it was necessary to acquire previously unavailable knowledge about the connectional neuroanatomy of the PMCs. For this reason, our research group undertook an experimental neuroanatomical study in nonhuman primates. The experiments were conducted in Josef Parvizi's laboratory in collaboration with Gary Van Hoesen. In essence the study consisted of making, in experimental macaque monkeys, numerous injections of biological tracers into all the territories whose neural connectivity we needed to investigate. Once injected into a given brain region, biological tracers are absorbed by individual neurons and transported along their axons all the way to their natural destinations, whatever the neurons are currently connecting to. These are the so-called anterograde tracers. Another kind of biological tracer, the retrograde kind, is taken up by axon terminals and transported in reverse, from wherever the terminals are, back to the cell bodies of the neurons, at their points of origin. The upshot of all the tracer travels is the possibility of charting, for each target region, the sites of origin of the connections the region receives, as well as the sites toward which the region sends its messages.

The PMCs are constituted by several subregions. (In Brodmann's cytoarchitectonic map, they are areas 23a/b, 29, 30, 31, and 7m.) The interconnectivity of these subregions is so intricate that it is reasonable, to some degree, to treat them as a functional unit. Some distinct connectional affiliations within the subsectors open the possibility that some of them may have distinct functional roles to play. The umbrella term we coined for the ensemble appears justified, at least for the time being.



Figure 9.3: The location of the posteromedial cortices in the human brain.

The pattern of PMC connections, as reported in the first publication to come from these laborious and time-consuming investigations,³ is summarized in Figure 9.4. It can be described as follows:

1. Inputs from parietal and temporal association cortices, entorhinal cortices, and frontal cortices converge in the PMCs, as do inputs from the anterior cingulate cortex (a principal recipient of projections from the insula), the claustrum, the basal fore-brain, the amygdala, the premotor region, and the frontal eye fields. Thalamic nuclei, both intralaminar and dorsal, also project to the PMCs.

2. With few exceptions, the sites that originate converging inputs to the PMCs also receive diverging outputs from them, exceptions being the ventromedial prefrontal cortex, the claustrum, and the intralaminar nuclei of the thalamus. Some sites that do not project to the PMCs do receive PMC projections, namely the caudate and putamen, the nucleus accumbens, and the periaqueductal gray.

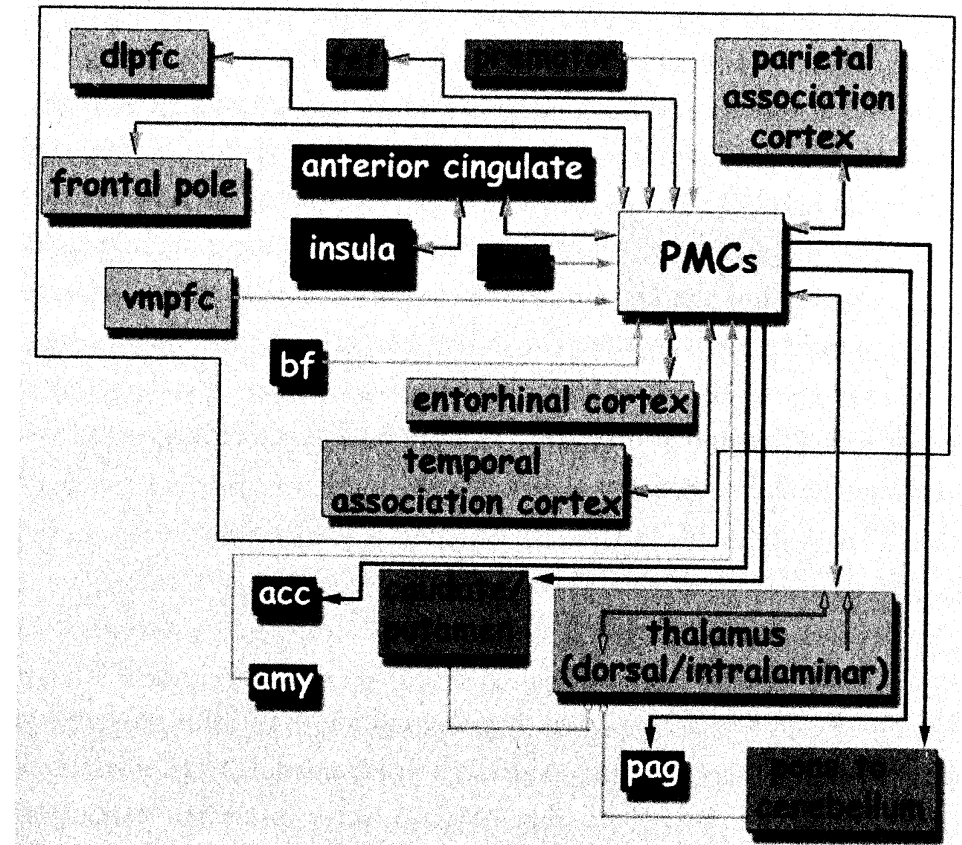


Figure 9.4: The pattern of neural connections to and from the posteromedial cortices (PMCs), as determined in a study conducted in the monkey. Abbreviations: dlpfc = dorsolateral prefrontal cortex; fef = frontal eye fields; vmpfc = ventromedial prefrontal cortex; bf = basal forebrain; claus = claustrum; acc = nucleus accumbens; amy = amygdala; pag = periaqueductal gray.

3. There are no connections to or from the PMCs relative to the early sensory cortices or the primary motor cortices.

4. From the results described under 1 and 2, it is apparent that the PMCs are a high-level convergence and divergence region. It is a prominent member of the club of CDRegions that I regard as good candidates for coordinating the contents in the conscious mind, and it even has an important connection with another potential coordinator, the claustrum, which significantly projects to the PMCs but is poorly reciprocated.

A recent study conducted in humans has added support for the idea that the PMCs are neuroanatomically distinct.⁴ The study, which was led by Olaf Sporns, used a modern technique of magnetic resonance imaging, diffusion spectrum imaging, which produces images of neural connections and of their approximate spatial distribution. The authors used their imaging data to construct maps of the connectional arrangements throughout the human cerebral cortex. They identified several connectional hubs throughout the cerebral cortex, several of which correspond to the CDRegions I have been discussing. They also concluded that the PMC region constitutes a unique hub, more strongly interrelated to other hubs than any of the others.

The PMCs at Work

We are now in a better position to imagine how the PMCs might contribute to the conscious mind. Although this is a sizable portion of cerebral cortex, the power of the PMCs resides not with territorial possessions but with the company they keep. The PMCs receive signals from most high-order sensory association regions and premotor regions and largely return the favors. Brain areas rich in convergence-divergence zones, which hold the key to composites of multimodal information, are thus able to signal to the PMCs and by and large can be signaled back. The PMCs also receive signals from subcortical nuclei involved in wakefulness and in turn signal to a variety of subcortical regions related to attention and reward (in the brain stem and basal forebrain), as well as to regions capable of producing motor routines (such as the basal ganglia and the periaqueductal gray).

What are the received signals likely to be about, and what do the PMCs do with them? We do not know for certain, but the huge disproportion between the profusion and strength of the projections toward the PMCs and the actual territory in which they land suggests an answer. The PMCs are mostly of older vintage, territories that one

thinks of as holding dispositions rather than explicit maps. The PMCs are not modern early sensory cortices like those of vision or hearing, where detailed maps of things and events can be assembled. Let us say that the PMC gallery has not enough wall space to exhibit large paintings or, for that matter, to present puppet shows. But that is just fine because the cortices that signal to the PMCs are not like early sensory cortices either; they cannot exhibit large paintings or present puppet shows any more than the PMCs can; they too are largely dispositional, convergence-divergence zone holders of recorded information.

Given their design, the PMCs as a whole and their component submodules are likely to behave as convergence-divergence regions themselves. I envision that the information held by the PMCs as well as by their partners can be played back only by signaling back into other CDRegions in the club, which in turn can signal to early sensory cortices. Those are the cortices where images can be made and displayed—that is, where large paintings can be shown and puppet shows presented. Relative to the other convergence-divergence regions that interconnect with them, the PMCs have a special hierarchical rank. The PMC region sits higher on the totem pole, capable of interactive signaling with the other CDRegions.

How, then, does the PMC assist consciousness? By contributing to the assembly of autobiographical self states. This is what I envision: separate sensory and motor activities related to personal experience would have been originally mapped in the appropriate brain regions, cortically and subcortically, and the data recorded in convergence-divergence zones and in convergence-divergence regions. In turn, the PMCs would have constituted a higher-order CDRegion record interconnected with the other CDRegions. The arrangement would allow activity in the PMCs to access larger, highly distributed data sets, but with the advantage that the access command would come from a relatively small and thus spatially manageable territory. The PMCs could support the establishment of momentary and temporally cohesive displays of knowledge.

If the PMCs' pattern of neuroanatomical connections is noteworthy, so is their anatomical location. The PMCs are located near the midline, the left set looking across the interhemispheric divide at the right set. This geographic position within the brain volume is convenient for both convergence and divergence connectivity relative to most regions of the cortical mantle, and it is ideal for receiving signals from the thalamus and reciprocating them. Curiously, the location also affords protection from external impact, and, because it is supplied by three major and separate blood vessels, it makes the PMCs relatively immune to the sort of vascular damage or trauma that could radically destroy them.

As I have previously emphasized, consciousness-related structures share several anatomical traits. First, either at the subcortical or the cortical level, they tend toward the old vintage. This should not be surprising given that the beginnings of consciousness occurred late in biological evolution but are not at all a recent evolutionary development. Second, both cortical and subcortical structures tend to be placed at or near the midline, and, just like the PMCs, they like to look at their twin siblings across the brain's midline—this is the case with thalamic and hypothalamic nuclei, as well as with brain-stem tegmental nuclei. Evolutionary age and convenience of location relative to widespread signal distribution are closely correlated here.

The PMCs would operate as a partner to the network of cortical CDRegions. But the role of the other CDRegions and the importance of the protoself system is such that consciousness is likely to be affected but not abolished following the hypothetical destruction of the entire PMC region, provided all the other CDRegions and the protoself system remain intact. Consciousness would be restored, albeit not at its peak. The situation of late-stage Alzheimer's disease, which I describe in the next section, is different in the sense that the PMC insult is virtually the last straw in a process of gradual ravage that has already disabled other CDRegions and the protoself system.

Other Considerations on the Posteromedial Cortices

ANESTHESIA RESEARCH

In some respects, general anesthesia is an ideal means to investigate the neurobiology of consciousness. It is one of the most spectacular developments of medicine and has saved the lives of millions of people who otherwise could not have had surgery. One often thinks of general anesthesia as a painkiller, since its effects preclude the pain that surgical wounds would cause, but the truth is that anesthesia precludes pain in the most radical way possible: it suspends consciousness altogether, not merely pain but all aspects of the conscious mind.

Superficial levels of anesthesia reduce consciousness lightly, leaving room for some unconscious learning and the occasional "breakthrough" of conscious processing. Deep levels of anesthesia cut deep into the conscious process and are, in point of fact, pharmacologically controlled variations on the vegetative state or even coma. That is what your surgeon needs if he is to work in peace inside your heart or your hip joint. You must be far, far away from it all, so deeply asleep that your muscular tone is as tough as jelly and you are not able to move. Stage III anesthesia is the ticket, and at that stage you will hear nothing, feel nothing, and think of nothing. When the surgeon talks to you, you will not respond.

The history of anesthesia has provided surgeons with numerous pharmacological agents to work with, and the search for the molecules that can do the most efficient job with minimal risks and little toxicity is an ongoing effort. By and large, anesthetics do their job by increasing inhibition in neural circuits. This can be achieved by strengthening the action of GABA (gamma-aminobutyric acid), the leading inhibitory transmitter in the brain. Anesthetics act by hyperpolarizing neurons and blocking acetylcholine, an important molecule in normal neuron-to-neuron communication. It was commonly thought that anesthetic

agents worked by depressing brain function across the board, bringing down the activity of neurons most everywhere. But recent studies have shown that some anesthetics work very selectively, exerting their action at specific brain sites. A case in point is propofol. As shown in functional neuroimaging studies, it does its splendid job by working principally at three sites: the posteromedial cortices, the thalamus, and the brain-stem tegmentum. While the relative importance of each site in the production of unconsciousness is unknown, the decreases in level of consciousness are correlated with the decrease of regional blood flow in the posteromedial cortices.⁵ But the evidence goes well beyond propofol. Other anesthetic agents seem to have comparable effects, as a comprehensive review demonstrates. Three paramedian brain territories instrumental in building consciousness are selectively depressed by propofol anesthesia.

SLEEP RESEARCH

Sleep is a natural setting for the study of consciousness, and sleep studies were early contributors to the understanding of the problem. It has been well established that electroencephalographic rhythms, the distinct patterns of electrical activity generated by the brain, are associated with specific stages of sleep. It is notoriously difficult to peg the origin of electroencephalographic patterns to particular brain regions, and that is where the spatial localization of functional neuroimaging techniques has come in handy to complete the picture. Using imaging techniques, it has been possible, over the past decade, to take a closer look at specific brain regions during varied stages of sleep.

For example, consciousness is deeply depressed during slow-wave sleep, also known as non-rapid eye movement sleep or N-REM. This is the deep slumber of the kind and the just, the slumber from which only the unkind and unjust alarm clock will wake us up. This is “dreamless sleep,” although the complete absence of dreams appears to apply only to the first part of the night. Functional neuroimaging studies show

that in slow-wave sleep, activity is reduced in a number of brain regions, most prominently in parts of the brain-stem tegmentum (at the pons and midbrain), the diencephalon (the thalamus and the hypothalamus/basal forebrain), the medial and lateral parts of the prefrontal cortex, the anterior cingulate cortex, the lateral parietal cortex, and the PMCs. The pattern of functional reduction in slow-wave sleep is less selective than in general anesthesia (there is no reason why the pattern should be the same), but as in anesthesia, it does not suggest an across-the-board depression of function. The pattern does include, prominently, the three correlates of consciousness-making (brain stem, thalamus, and PMCs), and it does show that all three are depressed.

Consciousness is also depressed during rapid eye movement (REM) sleep, during which dreams are most prevalent. But REM sleep allows dream contents to enter consciousness, either via learning and subsequent recall or via so-called paradoxical consciousness. The brain regions whose activity is most markedly decreased during REM are the dorsolateral prefrontal cortex and the lateral parietal cortex; predictably, the decrease in activity of the PMCs is far less marked.⁶

In brief, the level of activity in the PMCs is highest during wakefulness and lowest during slow-wave sleep. During REM sleep the PMCs operate at intermediate levels. This makes some sense. Consciousness is mostly suspended during slow-wave sleep; in dream sleep, things do happen to a “self.” The dream self is not the normal self, of course, but the brain state that goes with it appears to recruit the PMCs.

THE PMCS' INVOLVEMENT IN THE DEFAULT NETWORK

In a series of functional imaging studies using both positron-emission tomography and functional magnetic resonance, Marcus Raichle called attention to the fact that when subjects are at rest, not engaging in a task requiring focused attention, a selective subset of brain regions appears consistently active; when attention is directed to a specific task, the activity of these regions decreases slightly, but never to the degree

