Consciousness and Anesthesia

Michael T. Alkire, Anthony G. Hudetz, Giulio Tononi

When we are anesthetized, we expect consciousness to vanish. But does it always? Although anesthetics undoubtedly induce unresponsiveness and amnesia, the extent to which it causes unconsciousness is harder to establish. For instance, certain anesthetics act on areas of the brain’s cortex near the midline and abolish behavioral responsiveness, but not necessarily consciousness. Unconsciousness is likely to ensue when a complex of brain regions in the posterior parietal area is inactivated. Consciousness vanishes when anesthetics produce functional disconnection in this posterior complex, interrupting cortical communication and causing a loss of integration; or when they lead to bistable, stereotypic responses, causing a loss of information capacity. Thus, anesthetics seem to cause unconsciousness when they block the brain’s ability to integrate information.

How consciousness arises in the brain remains unknown. Yet, for nearly two centuries our ignorance has not hampered the use of general anesthesia for routinely extinguishing consciousness during surgery. Unfortunately, once in every 1000 to 2000 operations a patient may temporarily regain consciousness or even remain conscious during surgery (1). Such intraoperative awareness arises in part because our ability to evaluate levels of consciousness remains limited. Nevertheless, progress is being made in identifying general principles that underlie how anesthetics bring about unconsciousness (2–6) and how, occasionally, they may fail to do so.

Cellular Actions of Anesthetics

The cellular and molecular pharmacology of anesthetics has been reviewed extensively (6–8). General anesthetics fall into two main classes: intravenous agents used to induce anesthesia, generally administered together with sedatives or narcotics; and volatile agents, generally used for anesthesia maintenance (Table 1). Anesthetics are thought to work by interacting with ion channels that regulate synaptic transmission and membrane potentials in key regions of the brain and spinal cord. These ion-channel targets are differentially sensitive to various anesthetic agents (Table 1).

Anesthetics hyperpolarize neurons by increasing inhibition or decreasing excitation (9) and alter neuronal activity: The sustained firing typical of the aroused brain changes to a bistable burst-pause pattern (10) that is also observed in non–rapid-eye-movement (NREM) sleep. At intermediate anesthetic concentrations, neurons begin oscillating, roughly once a second, between a depolarized up-state and a hyperpolarized down-state (11). The up-state is similar to the sustained depolarization of wakefulness. The down-state shows complete cessation of synaptic activity for a tenth of a second or more, after which neurons revert to another up-state. As anesthetic doses increase, the up-state turns to a short burst and the down-state becomes progressively longer. These changes in neuronal firing patterns are reflected in the electroencephalogram (EEG) (electrical recording from the scalp) as a transition from the low-voltage, high-frequency pattern of wakefulness (known as activated EEG), to the slow-wave EEG of deep NREM sleep, and finally to an EEG burst-suppression pattern (12).

The Anesthetized Patient: Unconscious or Unresponsive?

Clinically, at low-sedative doses anesthetics cause a state similar to drunkenness, with analgesia, amnesia, distorted time perception, depersonalization, and increased sleepiness. At slightly higher doses, a patient fails to move in response to a command and is considered unconscious. This behavioral definition of unconsciousness, which was introduced with anesthesia over 160 years ago, while convenient, has drawbacks. For instance, unresponsiveness can occur without unconsciousness. When we dream, we have vivid conscious experiences, but are unresponsive because inhibition by the brainstem induces muscle paralysis (13). Similarly, paralyzing agents used to prevent unwanted movements during anesthesia do not remove consciousness (14).

Certain anesthetics may impair a person’s willfulness to respond by affecting brain regions where executive decisions are made. This is not an issue for anesthetics that globally deactivate the brain, but it may be problematic for dissociative anesthetics like ketamine. Low doses of ketamine cause depersonalization, out-of-body experiences, forgetfulness, and loss of motivation to follow commands (15). At higher doses, ketamine causes a characteristic state in which the eyes are open and the face takes on a disconnected blank stare. Neuroimaging data show a complex pattern of regional metabolic changes (16), including a deactivation of executive circuits in anterior cingulate cortex and basal ganglia (Fig. 1) (17). A similar open-eyed unresponsiveness is seen in akinetiform mutism after bilateral lesions around the anterior cingulate cortex (18). In at least some of these cases, patients understand questions, but may fail to respond. Indeed, a woman with large frontal lesions who was clinically unresponsive was asked to imagine playing tennis or to navigate her room, and she showed cortical activation patterns indistinguishable from those of healthy subjects (19). Thus, clinical unresponsiveness is not necessarily synonymous with unconsciousness.

At doses near the unconsciousness threshold, some anesthetics block working memory (20). Thus, patients may fail to respond because they immediately forget what to do. At much lower doses, anesthetics cause profound amnesia. Studies with the isolated forearm technique, in which a tourniquet is applied to the arm before paralysis is

---

1Department of Anesthesiology and the Center for the Neurobiology of Learning and Memory, University of California, Irvine, CA 92868, USA. 2Department of Anesthesiology, Medical College of Wisconsin, Milwaukee, WI 53226, USA. 3Department of Psychiatry, University of Wisconsin, Madison, WI 53719, USA.

*To whom correspondence should be addressed. E-mail: gtononi@wisc.edu

Table 1. Ionic mechanisms and targets of current clinical anesthetics (6, 8). Abbreviations: Ach, acetylcholine; AMPA, α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid; GABA_A, γ-aminobutyric acid, type A; NMDA, N-methyl-D-aspartate.
induced (to allow the hand to move while the rest of the body is paralyzed), show that patients under general anesthesia can sometimes carry on a conversation using hand signals, but postoperatively deny ever being awake (21). Thus, retrospective oblivion is no proof of unconsciousness.

Nevertheless, at some level of anesthesia between behavioral unresponsiveness and the induction of a flat EEG [indicating the cessation of the brain’s electrical activity, one of the criteria for brain death (22)], consciousness must vanish. Therefore, the use of brain-function monitors could improve consciousness assessment during anesthesia (23). For instance, bispectral index monitors record the EEG signal over the forehead and reduce the complex signal into a single number that tracks a patient’s depth of anesthesia over time (12). Such devices help guide anesthetic delivery and may reduce cases of intraoperative awareness (24), but they remain limited at directly indicating the presence or absence of consciousness, especially around the transition point. The isolated forearm technique has shown that individual patients can be aware and responsive during surgery even though their bispectral index value suggests they are not (25). Either the EEG is not sensitive enough to the neural processes underlying consciousness, or we still do not yet fully understand what to look for.

The Thalamus—Switch or Readout?
The most consistent regional effect produced by anesthetics at (or near) loss of consciousness is a reduction of thalamic metabolism and blood flow (Fig. 1), suggesting that the thalamus may serve as a consciousness switch (2). Indeed, switchlike effects have been found with a number of thalamic manipulations. For example, γ-aminobutyric acid (GABA) agonists (mimicking anesthetic action) injected into the intralaminar nuclei cause rats to rapidly fall asleep, with a corresponding slowing of the EEG (26). Conversely, rats under anesthetic concentrations of sevoflurane can be awakened by a minute injection of nicotine into the intralaminar thalamus (27). In humans, midline thalamic damage can result in a vegetative state (17). Conversely, recovery from the vegetative state is heralded by the restoration of functional connectivity between thalamus and cingulate cortex (28). Also, deep brain electrical stimulation of the central thalamus improved behavioral responsiveness in a patient who was minimally conscious (29).

Nevertheless, thalamic activity does not decrease with all anesthetics. Ketamine increases global metabolism, especially in the thalamus (16). Other anesthetics can substantially reduce thalamic activity at doses that cause sedation, not unconsciousness. For instance, sevoflurane sedation causes a 23% reduction of relative thalamic metabolism when subjects are still awake and responsive (30). Indeed, anesthetic effects on the thalamus may be largely indirect (31, 32). Spontaneous thalamic firing during anesthesia is largely driven by feedback from cortical neurons (33), especially anesthetic-sensitive layer V cells (34). Many of these cells also project onto brainstem arousal centers, so cortical deactivation can reduce both thalamic activity and arousal (35). Also, the metabolic and electrophysiological effects of anesthetics on the thalamus in animals are abolished by removal of the cortex (33, 34, 36). By contrast, after thalamic ablation, the cortex still produces an activated EEG (37), suggesting that the thalamus is not the sole mediator of cortical arousal, nor perhaps is it the most direct one. In patients with implanted brain electrodes undergoing a second surgery to place a deep brain stimulator, the cortical EEG changed dramatically the instant the patients lost consciousness (38). However, there was little change in thalamic EEG activity until 10 min later. Conversely, in epileptic patients, during REM sleep (usually associated with dreaming) the cortical EEG was activated as if patients were awake, but the thalamic EEG showed slow wave activity, as if patients were asleep (39). Thus, the effects of anesthetics on the thalamus may represent a readout of global cortical activity rather than a consciousness switch, and thalamic activity may not be a sufficient basis for consciousness.

Nonetheless, it is premature to write off the thalamus altogether. Perhaps efficient communication among cortical areas requires a thalamic relay (40), in which case thalamic lesions would lead to a functional disconnection despite an activated cortex. A functional thalamic disconnection during anesthesia has been found with neuroimaging (41). Subthreshold depolarization to many cortical areas may be provided by calbindin-positive matrix cells, which are especially concentrated within some intralaminar thalamic nuclei and project diffusely to superficial layers of cortex (42). Cells in intralaminar nuclei can fire at high frequencies, thus providing a coherent oscillatory bias that may facilitate long-range cortico-cortical interactions. Therefore, whereas cortical arousal may occur without the thalamus, consciousness may not.

Cortical Effects of Anesthetics
Are some cortical areas more important than others for the induction of unconsciousness by anesthetics? Evoked responses in primary sensory cortices—the first relay for incoming stimuli—are often unchanged during anesthesia, deep sleep, and in vegetative patients. Also, activity in primary sensory areas often does not correlate with perceptual experience (43). Frontal cortex too may not be essential for anesthetic unconsciousness, because different anesthetics have variable effects on this area. For instance, at equivalent hypnotic doses, both propofol and thiopental deactivate posterior brain areas, but only propofol deactivates frontal cortex (44). Furthermore, large lesions of the frontal cortex do not by themselves produce unconsciousness (45).

Anesthetic-induced unconsciousness is usually associated with deactivation of mesial parietal cortex, posterior cingulate cortex, and precuneus (Fig. 1) (46). These same areas are deactivated in vegetative patients but are the first to reactivate in those who recover (28). Moreover, neural activity in these areas is altered during seizures associated with an impairment of consciousness (47) and in sleep (48). These mesial cortical areas are strategically located at the main hub of the brain’s connection core (49). They are also part of a default network that is especially active at rest and may be involved in global monitoring of the internal environment and several functions related to the self (50). Nevertheless, mesial cortical areas are deactivated in REM sleep (48), when subjects experience vivid dreams. Intriguingly, at intermediate doses, certain anesthetics, such as nitrous oxide, produce a fairly selective deactivation of posterior mesial cortex (51), yet when these areas start to turn off, subjects report dreamlike feelings with depersonalization and out-of-body experiences, rather than unconsciousness.

In addition to mesial cortical areas, many anesthetics also deactivate or disconnect a lateral temporoparieto-occipital complex of multimodal associative areas centered on the inferior parietal cortex (Fig. 1). In this case, lesion and anesthesia data are mutually supportive: Patients with bilateral lesions at the temporoparieto-occipital junction show no sign of perceptual experience, despite a flurry of undirected motor activity, a condition called hyperkinetic mutism (18). Thus, a complex of posterior brain areas comprising the lateral temporoparieto-occipital junction and perhaps a mesial cortical core are most likely the final common target for anesthetic-induced unconsciousness.

Disruption of Cortical Integration
Loss of consciousness may not necessarily require that neurons in these posterior brain areas be inactivated. Instead, it may be sufficient that dynamic aspects of neural activity change, especially if these affect the brain’s ability to integrate information (Fig. 2) (3, 5).
Consider first large-scale integration, loosely defined as the ability of different cortical regions to interact effectively (52). When consciousness fades during anesthesia, there is a drop in EEG coherence in the \(\gamma\)-frequency range (20 to 80 Hz) between right and left frontal cortices and between frontal and occipital regions (4). Anesthetics also suppress fronto-occipital \(\gamma\) coherence in animals, both under visual stimulation and at rest (53). The effect is gradual and much stronger for long-range than for local coherence (53). Anesthetics may disrupt cortical integration (5) by acting on structures that facilitate long-range cortico-cortical interactions, such as the posterior cortical connectional hub (49), certain thalamic nuclei (42), or possibly the claustrum (54). Anesthetics may also disrupt synchronization among distant areas by slowing neural responses (55).

The loss of feedback interactions in the cortex may be especially critical. When rats become unresponsive under anesthesia, information transfer first decreases in the feedback direction (Fig. 2B) (56). Also, anesthesia suppresses the late component (>100 ms) of visual responses, possibly by inhibiting feedback connections (57), but not the early feedforward components. Moreover, anesthesia abolishes contextual and attentional modulation of firing, presumably mediated by feedback connections (58). The corticothalamic system may be especially vulnerable to anesthetics due to its small-world organization. Small-world networks have mostly local connectivity with comparatively few long-range connections. Augmented with hubs, such networks maximize interactions while minimizing wiring. By the same token, anesthetics need only disrupt a few long-range connections to produce a set of disconnected components. Indeed, computer simulations demonstrate a rapid state transition at a critical anesthetic dose (59), consistent with a breakdown in network integration.

Disruption of Cortical Information Capacity
Consider next how anesthetics affect information, defined loosely as the number of discriminable activity patterns. When the repertoire of discriminable firing patterns available to the corticothalamic system shrinks, neural activity becomes less informative, even though it may be globally integrated (52). As described above, at high enough doses several anesthetics produce a burst-suppression pattern in which a near-flat EEG is interrupted every few seconds by brief, quasi-periodic bursts of global activation—a stereotypic, global on-off pattern. Such stereotypic burst-suppression can also be elicited by visual, auditory, and mechanical stimuli (Fig. 3B) (60, 61). Thus, during deep anesthetic unconsciousness, the corticothalamic system can still be active—in fact, hyperexcitable—and can produce global responses. However, the repertoire of responses has shrunk to a stereotypic burst-suppression pattern, with a corresponding loss of information, essentially creating a system having only two possible states (on or off). Generalized convulsive seizures provide another example in which consciousness can be lost even though neural activity remains high and highly

Fig. 2. Unconsciousness is associated with a loss of cortical integration. (A) The corticothalamic system is represented metaphorically as a large die having many faces, each corresponding to a different brain firing pattern. During conscious waking, the die rolls on a particular face, ruling out all the others and thus generating integrated information. If integration is lost (as in anesthesia or sleep), the die disintegrates into many two-faced dice, each generating 1 bit of information. (B) Anesthesia reduces cortical integration in the rat. (Top) During waking, transfer entropy, a measure of directional interactions among brain areas, is balanced in the feedforward (green) and feedback (red) directions. During anesthesia, feedback transfer entropy (red) is reduced, implying a decrease in front-to-back interactions. (Bottom) Responses to a flashing light delivered at 0.2 Hz (arrow) from a representative rat when awake and under 1.1% isoflurane anesthesia (56). When the rat is awake, each flash evokes a sustained \(\gamma\)-frequency (20 to 60 Hz) response in visual occipital cortex (blue) and a later response in parietal association cortex (red). During anesthesia, the occipital response is preserved, although it is shorter (blue), and the parietal response is attenuated, indicating that anesthesia reduces cortical interactions and thus reduces integration. (C) Sleeping reduces cortical integration in humans. EEG voltages and current densities are shown from a representative subject in which the premotor cortex was stimulated with transcranial magnetic stimulation (TMS) (black arrow). During waking (top), stimulation evokes EEG responses first near the stimulation site (black circle; the white cross is the site of maximum evoked current) and then in sequence at other cortical locations. During deep sleep (bottom), the stimulus-evoked response remains local, indicating a loss of cortical integration.
synchronized: A large portion of the corticothalamic complex is engaged in strong, hypersynchronous activity, but this activity is stereotypic (60, 61).

A Bit Like Sleep
Sleep is the only time when healthy humans regularly lose consciousness. Subjects awakened during slow wave sleep early in the night may report short, thoughtlike fragments of experience, or often nothing at all (13). Although anesthesia is not the same as natural sleep, brain-arousal systems are similarly deactivated (6, 62). Also, as under anesthesia, during slow wave sleep, cortical and thalamic neurons become bistable and undergo slow oscillations (1 Hz or less) between up- and down-states. Like animal studies during anesthesia, during slow wave sleep, cortical and thalamic neurons become bistable and undergo slow oscillations (1 Hz or less) between up- and down-states. Like animal studies during anesthesia, the bistability of cortical neurons has consequences for the brain’s capacity to integrate information (Figs. 2C and 3C). During waking, transcranial magnetic stimulation (TMS) applied to premotor cortex and other cortical areas induces a sustained response (300 ms) involving the sequential activation of specific brain areas, the identity of which depends upon the precise site of stimulation (63, 64). During early NREM sleep, possibly due to the induction of a local down-state, TMS pulses produce instead a short (<150 ms) local response (64), suggesting a loss of integration. Intriguingly, TMS pulses to mesial parietal regions, overlying the main hub in the cortical connectional core (49), trigger a stereotypic, high-amplitude slow wave closely resembling spontaneous slow waves (63). This stereotypic response, presumably due to the simultaneous activation of the cortical connectional core and to the induction of a global down-state, reflects a limited repertoire of activity patterns and thus a loss of information.

Consciousness and Integrated Information
The evidence from anesthesia and sleep states (Fig. 2 and 3) converges to suggest that loss of consciousness is associated with a breakdown of cortical connectivity and thus of integration, or with a collapse of the repertoire of cortical activity patterns and thus of information (Figs. 2 and 3). Why should this be the case? A recent theory suggests a principled reason: Information and integration may be the very essence of consciousness (52). Classically, information is the reduction of uncertainty among alternatives: When a coin falls on one of its two sides, it provides 1 bit of information, whereas a die falling on one of six faces provides ~2.6 bits. But then having any conscious experience, even one of pure darkness, must be extraordinarily informative, because we could have had countless other experiences instead (think of all the frames of every possible movie). Having any experience is like throwing a die with a trillion faces and identifying which number came up (Figs. 2A and 3A). On the other hand, every experience is an integrated whole that cannot be subdivided into independent components. For example, with an intact brain you cannot

Fig. 3. Unconsciousness is associated with a loss of information capacity. (A) As in Fig 2, the corticothalamic system is represented metaphorically as a large die having many faces, each corresponding to a different brain firing pattern. During conscious waking, the die rolls on a particular face, ruling out all the others and thus generating integrated information. If information is lost (as in anesthesia or sleeping), the die is flattened so that it has only two faces (firing patterns). Due to the loss of repertoire, it generates only 1 bit of information. (B) Anesthesia reduces information capacity in rat cortex. (Top) Field potentials recorded before and during light flashes (marks below each trace). During waking (left), flash-evoked field potentials (blue) (light flashes indicated by marks below each trace) are small and variable, being masked by spontaneous neuronal activity. During deep anesthesia (right), bursts of activity occur spontaneously and after each light flash. (Bottom) During anesthesia, the γ-burst response is uniform across all three brain regions. Thus, responses are stereotypic and lack regional specificity, indicating a loss of information capacity. (C) Sleeping reduces cortical information carrying capacity in humans. (Top) During waking, stimulation over the mesial parietal cortex produces a specific, sequential pattern of activation. (Bottom) During sleep, stimulation produces a global, stereotypic response that spreads from the stimulation site to most of the cortex, indicating a loss of information capacity. Black traces represent averaged voltage potentials recorded at all electrodes and superimposed, while estimated current density is displayed in absolute scale (63, 64).
experience the left half of the visual field independently of the right half, or visual shapes independently of their color. In other words, the die of experience is a single one—throwing multiple dice and combining the numbers will not do.

Less metaphorically, the theory claims that the level of consciousness of a physical system is related to the repertoire of different states (information) that can be discriminated by the system as a whole (integration). A measure of integrated information, called \( \Phi \), can be used to quantify the information generated when a system enters one particular state of its repertoire, above and beyond the information generated independently by its parts (52, 65). In practice, \( \Phi \) can only be measured rigorously for small, simulated systems. However, empirical measures could be devised to evaluate integrated information on the basis of EEG data, resting functional connectivity, or TMS-evoked responses. This approach could allow the development of consciousness monitors that evaluate both loss of integration, as revealed by reduced functional or effective connectivity, and loss of information, as evidenced by stereotypical responses.

This theory has some interesting implications for anesthesia. For example, it explains why a corticothalamic complex is essential for consciousness and is thus the proper target for anesthesia: By conjointing functional specialization (each cortical area and neuronal group within each area is exquisitely specialized) with functional integration (thanks to extensive corticocortical and corticothalamic connectivity), a corticothalamic complex is well suited to behave as a single dyadic entity endowed with a large number of discriminable states. By contrast, parts of the brain comprising the medial cortical core. Whether the medial or lateral component is more important, and whether anterior cortical regions are critical primarily for executive functions and perhaps self-reflection, remain questions for future work. Second, anesthetics can cause unconsciousness not just by deactivating this posterior corticothalamic complex, but also by producing a functional disconnection between subregions of this complex. Third, although assessing loss of consciousness with verbal commands may usually be adequate, it may occasionally be misleading. Finally, one theoretical framework that seems to fit well with current empirical data suggests that consciousness requires an integrated system with a large repertoire of discriminable states. According to this framework, anesthetics would produce unconsciousness either by preventing integration (blocking the interactions among specialized brain regions) or by reducing information (shrinking the number of activity patterns available to cortical networks). Other frameworks for consciousness, emphasizing access to a global workspace (66, 67), or the formation of large coalitions of neurons (43), are also consistent with many of the findings described here, especially those concerning the role of cortical integration. Together, these ideas should help in developing agents with more specific actions, in better monitoring their effects on consciousness, and in using anesthesia as a tool for characterizing the neural substrates of consciousness.

Conclusions

Despite different mechanisms and sites of action, most anesthetic agents appear to cause unconsciousness by targeting, directly or indirectly, a posterior lateral corticothalamic complex centered around the inferior parietal lobe, and perhaps a medial cortical core. Whether the median or lateral

**References and Notes**