# 2 Consciousness and the Freedom to Act

*Merlin Donald*

> "The only states of consciousness that we naturally deal with are found in personal consciousnesses, minds, selves, concrete particular I's and you's. Each of these minds keeps its own thought to itself. There is no giving or bartering between them. No thought ever comes into direct *sight* of a thought in another personal consciousness than its own. Absolute insulation, irreducible pluralism, is the law. It seems as if the elementary psychic fact were not *thought* or *this thought* or *that thought*, but *my thought*, every thought being *owned*."
>
> —William James, *Principles*.

The denial of personal efficacy is an old theme in Western civilization. Quite apart from the concept of fate, which dominated the ancient world, much post-Enlightenment thought has viewed scientific determinism as axiomatic. La Mettrie wrote *L'Homme Machine* almost three centuries ago. Skinner wrote *Beyond Freedom and Dignity* almost 40 years ago. Recently the attack on the concept of free will has resumed, and while earlier claims were argued mostly from questionable a priori principles, the new ones are supposedly supported by empirical research.

A number of modern cognitive scientists have argued that consciousness is merely a comforting illusion. Recently, volition and voluntary movement

have been relegated to the
although we may have the
make a decision to act in a
evidence. Even so small a
outside our conscious contr
The illusion of having "cause
in memory, and may indeed

This idea has a preceden
nated the first half of the
deterministic principles clai
that we are unconscious
Brain activity, inaccessible to
of any kind, including volun
preceded by an inaccessible
Our awareness of such a deci
make; the brain has already
informed of this event after
are constantly being shaped
permeate the background of
we make. There is nothing
periodic shift in the *Zeitgeist*
is shaped unconsciously, and

The empirical evidence
two findings: the timing of
evidence of unconscious inf
electrical activity are often
the efficacy of consciousness
argument. The claim is that
therefore cannot play a caus
concern this evidence, and its

The other argument, th
influences on awareness, is
because it is an inductive arg
scious influence. The most co
literature on priming, where
in perception and behavior in
outside the subject's awarene
Social psychology courses a
problem is that they only sho
scious influence; they do not
possibility of freedom of choic

Another well-known exp
influence concerns the aut

have been relegated to the same status (Wegner, 2002). The claim is that, although we may have the illusion of being free and fully aware when we make a decision to act in a certain way, this is contradicted by empirical evidence. Even so small a matter as a decision to move a finger is said to be outside our conscious control, determined by unconscious brain processes. The illusion of having "caused" the action is useful for us as a somatic marker in memory, and may indeed be comforting, but it is, after all, an illusion.

This idea has a precedent in the behavioristic movement that dominated the first half of the twentieth century. This crude application of deterministic principles claims to avoid dualism. The central argument is that we are unconscious automatons, and worse, deluded automatons. Brain activity, inaccessible to conscious awareness, always precedes action of any kind, including voluntary action. Even a decision to act must be preceded by an inaccessible brain signal, and therefore we are not truly free. Our awareness of such a decision comes too late to influence the choice we make; the brain has already initiated the action, and consciousness is only informed of this event after the fact. Moreover, our thoughts and behavior are constantly being shaped by unconscious factors. Unseen influences permeate the background of behavior and determine many of the choices we make. There is nothing new in this idea. Perhaps it just reflects a periodic shift in the *Zeitgeist* back to a new variant of positivism. Behavior is shaped unconsciously, and the unconscious is, by definition, not free.

The empirical evidence cited to defend this idea comes mainly from two findings: the timing of brain signals, relative to action, and strong evidence of unconscious influences on cognition. Recordings of brain electrical activity are often cited as conclusive negative evidence against the efficacy of consciousness in action. This is the heart of the timing argument. The claim is that consciousness always lags behind action and therefore cannot play a causal role in guiding it. The next section will concern this evidence, and its interpretation.

The other argument, that there are many powerful unconscious influences on awareness, is not fatal for free choice in its own right because it is an inductive argument. There are many instances of unconscious influence. The most convincing and elegant are found in the recent literature on priming, where countless experiments show subtle changes in perception and behavior induced by exposure to prior stimuli that are outside the subject's awareness at the time (Bargh & Ferguson, 2000). Social psychology courses are now full of such demonstrations. The problem is that they only show the ubiquity and importance of unconscious influence; they do not, and cannot in themselves, disprove the possibility of freedom of choice, or free will.

Another well-known experimental demonstration of unconscious influence concerns the automatization of performance. It is well established that highly skilled tasks are best performed when overpracticed, or automatized. If a subject directs second-by-second attention to skilled performance, the result can be disastrous. A pianist or gymnast who becomes self-conscious during a performance is less likely to do well. It is better to let the unconscious do the work. Similarly, creative thoughts often come in an unplanned, apparently unguided (therefore unconscious) free-associative manner. And so on. Priming effects and free-associative effects outside of awareness have been standard fare in psychology for a half a century or more—but there is a current thrust to overgeneralize such unconscious influences, so that they become the explanation for every aspect of thought and behavior, and no role is left for consciousness.

This type of demonstration is still inductive: it leaves the door open to some efficacious role for awareness. In order to rule out free will altogether, there must be convincing negative evidence that can be generalized, so that every kind of behavior and cognition has been accounted for, and consciousness has been left with no apparent function, and no empirical foundation. For some, that evidence has come from the timing argument.

## THE TIMING ARGUMENT

The apparent freedom entailed in voluntary action has been questioned on the basis that there is an inherent temporal distance between awareness and the brain processes that create it. In other words, awareness comes after the fact, and too late to influence choice. The most popular example used to support this idea consists of well-known experiments on finger movement carried out by Libet (1985, 1993). These experiments claim to show that brain activity "decides" to initiate a movement before the observer even becomes aware of having made a decision to move. If this were the case, it would constitute a possible death knell for the notion of conscious efficacy. Therefore it is important to examine these studies carefully.

The evidence for this claim is that a build-up of electrical activation in the premotor cortex occurs several seconds before the observer decides to initiate a movement. The technical name for the electrical wave that Libet recorded was variously labeled the "readiness potential" or "*Berietschaftspotential*" (BP) by its discoverers (Kornhuber & Deecke, 1965); another version of this was also known as the "expectancy wave" (Walter, Cooper, Aldridge, McCallum, & Winter, 1964). The activation takes the form of a surface-negative wave that peaks just before the action is performed and discharges just as the act is initiated. This effect has been replicated many times, in many laboratories, with various kinds of movements and electrical recordings. In my own laboratory, for instance, we studied this effect in the context of foot, hand, and mouth movement (Brooker & Donald, 1980; Donald, Inglis, Sproule, Young, & Monga, 1980). The results are similar in all three modalities of action. In a self-paced movement paradigm, when the subject freely decides to move, a surface-negative activation wave often begins well over a second before muscle contraction is initiated. It is true that it takes some time, in this case much less than 100 milliseconds, to send a motor command from the cortex to the hand muscles. It is the additional residual delay, between the onset of a BP and the movement, that must be explained.

Libet's special contribution to this paradigm was to add a subjective judgment to the self-paced movement task. The actors performing the movement were required to report the precise time when they decided to initiate the action, by reading the position of a moving second hand on a clock face. This subjective report was compared to the timing of muscle contraction, and the actual physical movement of the hand, as well as BP onset. Invariably, the subjects reported their decision to move long *after* the preparatory electrical activity had already begun. The delay between brain activity and awareness was consistently measured as three-quarters of a second or more. In parallel experiments on perception, Libet reported a time lag of approximately one-third of a second between electrocortical response to a cutaneous stimulus, and the time of subjective awareness of that stimulus. In both cases, conscious experience lagged behind brain activation.

Thus, it seems that consciousness only "catches up" with an action or a stimulus after it has already been processed thoroughly by the brain. Many theorists have used this result to infer that unconscious brain processes had already made the decision and started the movement, before subjects experienced the illusion that they had consciously decided to move. The implication here is that, in both sensation and action, awareness has no say in what happens. Awareness seemed to follow a physical brain event and always trailed behind the action. A more general proposition often follows: consciousness is inherently too late, after the fact, and powerless to influence behavior. Since there is no freedom without a conscious decision, there is no such thing as true freedom of choice.

These experiments have been criticized many times, and some of my own objections to this interpretation of Libet's work have appeared in print elsewhere (Donald, 2001). I have suggested that Libet's paradigm overemphasizes the sensorimotor interface and ignores the slower time range within which human conscious awareness matters most, within which executive and supervisory operations such as planning, metacognitive oversight, and social communication take place. I have also

> If you are not conscious of the decision the decision did not occur

suggested that we do not yet know enough about the slower-acting brain mechanisms that mediate such operations (Donald, 2008) to make pronouncements about the neural basis of consciousness. I will not repeat these points at length here.

However, there is another, more fundamental criticism of Libet's movement paradigm. In it, subjects are sitting in a room with their hands on a platform, and the only thing they have to do is move a finger, without moving anything else (not easy). For subjects trapped in such a situation, this decision to move becomes the sole focus of attention. It is an excruciatingly boring task, and it is not easy to sustain interest in it; as a result, there is much subjective anticipation of each decision to move.

The key fact here is that this kind of anticipation is very general. It is not really specific to the hand movement. The BP that it generates is not actually a motor command signal. Rather, it is a generalized preparatory wave that precedes any major self-initiated cognitive event. Libet's studies showed only that *the making of such a decision* was anticipated by the observer's brain long before the decision was made. This is hardly surprising, given the nature of the task, and the properties of the BP. It is also completely irrelevant to the question of conscious efficacy or free choice in moving the finger. While it is true that a BP appears several seconds before the action, the conclusion drawn in most studies of this effect prior to Libet's unusual claim, and by most investigators afterward, was that the BP was generated in anticipation of making a decision to move, rather than of moving per se.

In the commentaries on his 1985 BBS article, Libet had already been criticized for mistakenly judging that his subjects actually initiated a movement when the BP started to develop. They had not. Libet apparently confused the slow-moving BP with a later, last-minute spike known as the "command potential," which precedes a specific voluntary hand movement by approximately 100 ms and has both a different wave shape and a different cortical origin (Kornhuber & Deecke, 1965). The BP has been shown to be a sign of a very general state of anticipation. The fact that this mistake is still in circulation provides us with a very good example of how a misinterpretation of scientific evidence can be perpetuated and duplicated widely if the error matches the current expectations of scientists. The truth is, the timing argument fails on all counts.

## Monism, Dualism, and Psychophysical Simultaneity

But the conceptual problem goes deeper. However we may regard the meaning or validity of Libet's work, there is a more important challenge that haunts this field of research: the simple fact that brain activity is inaccessible to awareness. The implications of this fact have not been fully accepted or digested by many investigators.

Conscious observers can never be directly aware of the brain activity on which their awareness depends. One reason for this lies in simple anatomy. We are only able to sense things for which we have sensory receptors, and there are no specialized receptors to detect and record brain activity. Thus, there are no neural projection systems that could possibly inform some receptive region of the brain about its own electrical or chemical activity. Even if there were such receptors, their inputs would return home where they started, in another flurry of cortical activity, in an endless (and pointless) loop.

Surely this is a clear principle on which most scientists and philosophical schools of thought can agree: Brain activity as such cannot be directly experienced by the owner of that brain. While conscious experience has access to the *cognitive contents* of certain classes of brain activity, it cannot access the electrochemical activity itself. The subject experiences only the phenomenal outcome of the brain's activity. The brain's actions themselves—the firing of neurons, multidimensional changes in synaptic activity, exchanges of transmitter substances, changes in local current flow, and so on—are permanently out of bounds.

*Brain activity can never be the direct object of conscious experience.* This principle is basic, and not on the table. It is equally obvious, especially when examining electrical activity during epileptic seizures and observing the close temporal correlation between experience and behavior, that conscious experience is very much contingent on brain activity, and closely tethered to it in real time (because of this, cortical electrical activity was once the primary neurological criterion of brain death).

Thus, brain activity, while not consciously accessible, is essential for any kind of mental activity. A subset of that activity, as yet not properly identified, is dedicated to conscious mental activity. Consciousness thus seems to be produced by something that it cannot observe. Does this negate freedom of choice, reducing it to the status of an illusion? Does the mere fact that brain activity goes on outside of our awareness mean that we are automatons that live with the pleasant delusion that we are free?

These questions lead us back to a discussion of two old philosophical workhorses: monism and dualism. Monism is the position that most scientists claim to support in principle. It asserts the unity of science and denies dualism, which holds that body and mind are separate ontological domains. A monist stance on consciousness must therefore assert that brain activation and experience reflect a single underlying reality. Every experience must therefore have a corresponding brain state. No neuroscientist will question this.

If a given brain state and its corresponding experience are aspects of one and the same event, viewed from different vantage points, *there can be no direction of causality involved*. Neither aspect could precede the other in the causal chain leading to awareness—they are one and the same link in that chain. Nor could there be a temporal lag between brain and phenomenal experience. A specific brain state could not "precede," "lag," "trigger," or "lead toward" its corresponding subjective aspect. This has consequences for experimental design. An experience does not, could not, and should not be expected to follow or precede its corresponding brain aspect, as if it were crossing some mysterious ethereal divide. The two aspects, physical and phenomenological, must be simultaneous and identical.

Any other position amounts to dualism. This exposes the dualism in Libet's interpretation of his timing experiments. In a monist world, if an experimenter does not find simultaneity of brain activity and experience, he is looking in the wrong place!

There are many theories of consciousness compatible with a strict monist view. These include the "global workspace" hypothesis of Baars (1988) and Newman & Baars (1993), the "dynamic core" hypothesis of Edelman & Tononi (2000), and several theories of binding, which, while preliminary, provide hints of a possible physical correlate of consciousness. They also include some wider theories, such as my own (Donald, 2001), which try to unite the neural and cultural sources of awareness into a single theoretical framework.

This returns the discussion to the significance of studies like Libet's. Why would it matter that an expectancy wave was generated long before a decision? In a monist framework, such a finding could only imply that the real neural event corresponding to the experience of making a decision had not been found in that particular experiment, and that the experimenters should continue their search for the neural events related to the phenomenon under investigation. Any alternative conclusion would be dualistic.

As we move forward in building a neurocognitive theory of conscious choice, these issues will continue to come to the fore. Dualism should always be taken off the table in advance. Adherence to a monist approach to consciousness at our current stage of knowledge demands optimism and persistence. At present, we have not yet made enough progress even to know what the important neural variables are. But there is no reason to reject the possibility of developing a successful theory.

## CULTURE AND FREEDOM OF CHOICE

There is no evidence that would enable us to dismiss the possibility of free will and conscious choice. However, this fact does not result in an explanatory theory of any kind, nor does it resolve the issue, in the sense of "proving" that choices are truly free. It merely leaves the question open. Awareness, when viewed in terms of its physical or phenomenal aspect, must meet certain criteria before it can be called free. It is far from clear how a nervous system; that is, a physical system subject to the laws of physics and chemistry, can act in a manner that we would call free. It is equally uncertain how to apply a reasonable criterion of "freedomness" to a subjective experience.

But there are strong reasons to adhere to an evolutionary approach to consciousness. Consciousness exists in some form in many animals, and the expansion of the powers of conscious processing in evolution is mostly about improving the capacity for learning, especially the capacity to integrate complex new material that must be assimilated by the brain. The ultimate product of wider conscious integration (this principle applies to many species) is an expanded "world model" by which an organism can guide action, including its own future cognitive activity (thought is always potentially manifest in action). By implication, freedom of choice is linked to the breadth of an organism's worldview. The reach of consciousness, and freedom, is graded and flexible.

An evolutionary approach inevitably results in a graded theory of choice. As the individual mind's model of the world expands to include more and more possibilities and variables, so does freedom of choice. The window of awareness is narrow in simple organisms, and so choice is very restricted, perhaps limited to such things as control over emotion, or inhibition of response. This situation becomes wider and deeper with the emergence of more complex nervous systems, which must be equipped to deal with many more options and variables in determining any course of action.

The assessment of the width of awareness in animals is based on several variables, including sensitivity, curiosity, and the duration of working memory. The latter variable draws on Hebb's (1963) notion that what we experience as awareness is really the memory of immediate past experience, and depends on short-term memory. As the temporal integration of events in the world becomes more advanced, encompassing larger events, especially social events such as changing alliances, or shifts in the dominance hierarchy, awareness is confronted with many more possibilities. Choice must cut through this incredible complexity and resolve an adaptive pattern of action, and it must do this quickly and efficaciously.

Culture complicates this picture, especially in the case of humans. Humanity's uniqueness lies in our collective cognitive systems, that is, in our cultures. We have mind-sharing cultures, in which alliances of trust can be based on such cognitive things as a shared belief in the same ideas.

A simpler way to put this is that, as the hominid brain widened the range of experiences it could integrate in a single act of synthesis, it eventually reached a point where the collective interaction of such brains triggered a cultural chemistry that fed back onto the individual mind and brain. Cognitive-cultural interactions amplified the powers of individuals, who in turn amplified the collective powers of society, and so on, in what cyberneticists might have called a positive evolutionary feedback loop.

This feedback process stepped through a series of cognitive-cultural stages (cf. Donald, 1991), building more and more abstract networks of cognitive activity distributed across many brains, and eventually, across a hybrid system that included artificial memory technologies, and external symbolic devices, such as the cinema, that have effectively re-engineered experience itself.

The key adaptation enabling the emergence of human mind-sharing cultures, and eventually language itself, was an evolutionary expansion of the range of primate working memory. This expansion involved a qualitatively new neural "slow process" that was unique to social mammals and became very highly evolved in human beings. This process is constantly running in the background of awareness, while traditional mechanisms of attention, perception, and short-term memory run in the foreground. It is the ultimate governor of human mental activity.

The neural mediator of this slow process is still unknown, but such a process is a theoretical requirement to explain the cognitive properties of human culture (Donald, 2008). It operates in the very long time range of minutes and hours, and none of the known active electrochemical processes that have been related to cognition can operate in that time zone. Among the known properties of neurons, there are many possible candidates for such a process, but much work remains to be done in this area before it will be understood.

Apes share a similar cerebral architecture with humans but cannot master more than a fraction of human culture. This could be because their slow process is less fully evolved. There are very large differences between humans and apes in the complexity and temporal extent of the kinds of social events they can understand, and this is a crucial variable in determining the degree to which culture can be penetrated. Moreover, humans are much more accomplished learners, and much of that increased learning capacity depends upon slower-acting executive brain processes identified with conscious cognitive control.

In the case of humans, the claim of free will is based on special evidence and usually conveys the human subject's personal power to reflect upon, and choose, the direction of action, especially in the long term. We might agree that human beings do not normally wage war or build skyscrapers unconsciously! We plan, execute, and design such major adventures in groups, and do so very deliberately. There is no question that we also have the option to veto many choices that have already been made. While there are many, including myself, who defend the existence of some free will in animals, these claims are typically much less confident than in the human case and are usually expressed in terms of approximating, in varying degrees, human freedom of choice.

It is useful to treat the human case as special because it is the template, the standard to which others species are held. What is it that makes the human mind unique, and human choice uniquely free? The answer is that human uniqueness is founded in *the depth of our enculturation*. Culture has set the brain free in various ways, and allowed it to redeploy its resources. One major role for culture is found in language. Languages are inventions of the group; that is, they emerge from distributed cognitive networks. There are no documented cases in which isolated human individuals have invented languages or symbols in any form. Individuals have an innate potential for language, but it cannot be realized except in an interactive group. Therefore, inasmuch as language makes us more capable of free choice, this freedom must be attributed in part to collective cognitive processes, that is, culture, including its major cognitive offshoot, language.

Does language acquisition affect awareness? This is one question that can be answered affirmatively, and definitively, on the basis of empirical evidence. This effect can be seen most clearly in deaf people who acquired sign language as adults. Once they have language, they experience reality very differently and have great difficulty remembering anything of their previous lives. Helen Keller referred to her prelanguage self as "Phantom," and her experience has been corroborated in several other late language learners studied by Schaller (1991). It is possible that, to some degree, these subjects were capable of free choice before they had language, but there is no question that their experience of feeling free after learning language was directly related to their sense of empowerment over consciousness itself, especially over the power to consciously recall events and other items from memory. This power is affected by having language.

Moreover, autobiographical memory depends upon language, or more precisely, on the highest language capacity, the ability to construct a personal narrative. Neither narrative ability nor autobiographical memory can develop without early immersion in human culture (Nelson, 1993, 1996). Our sense of having freedom of choice, while not inconceivable outside the wide temporal framework of autobiographical memory, is apparently greatly constrained without it.

Another enormous source of cultural influence is found in the effect of symbolic memory technologies, including such things as writing and mathematical notations, on conscious awareness. Written displays reorganize working memory in the reader, who becomes plugged into a temporary distributed system. This distributed system reallocates the cognitive load imposed by various tasks, amplifying the parameters of working memory, and the powers of choice we experience. This is a direct function of having representations in artificial memory media (Donald, 2001). External displays leave the observer freer to reflect on representations in a way that is not available to natural, or biological, working memory. This is achieved by reorganizing the task, and offloading some operations to external media. The acquisition of literacy skills imposes a novel functional architecture on the brain. It actually reorganizes the way neural tissue is deployed in cognitive tasks (Donald, in press). In a variety of ways, symbol-using cultures liberate the brain, creating options that extend freedom of choice. Daniel Dennett (2003) has reached a somewhat similar conclusion about the importance of culture in evolving freedom of choice.

On the representational level, human awareness seems to be unique because of the effects of symbolization on the stream of consciousness. How could a capacity for symbolically mediated awareness have evolved in a brain that remains nonsymbolic in its internal workings? I proposed long ago that the symbolic process is imposed by culture, and that it evolved from the "outside in," consistent with a Vygotskian approach to language phylogenesis, as well as ontogenesis. In essence, the developing human brain is programmed by cultural immersion, in a process that has evolved through several successive stages. As this process widened the reach of awareness, it also increased the brain's ability to plan behavior, control attention, and maintain longer-term control over cognition itself.

## IMPLICATIONS FOR HUMAN FREEDOM OF CHOICE

The act of making a decision might be taken as the one major paradigm in which we might be able to examine a wide range of brain–culture interactions. Decisions are the final resolution events of a variety of cognitive scenarios that can engage, in theory, the entire voluntary action repertoire of human beings. Essentially, decisions occur at choice points in a cognitive sequence. Thus, one could, in theory, "decide" which memory to retrieve, which stimulus to attend to, whether to inhibit an impulse, which perceptual set to activate, which emotional attitude to assume toward a social scenario, and which pattern of action to pursue in a variety of contexts. Choice occurs everywhere in the higher cognitive system.

Choice engages most subroutines and subsidiary systems concerned with memory, symbolic representation, and thought. Thus, it constitutes an abstract category that cuts across all of cognition. Decisions are sometimes imposed from above on innate emotional brain systems—for instance, whether to release a smile in a particular social context—and are sometimes embedded in algorithms and mental habits with purely cultural origins—for instance, decisions involving numbers or other symbolic systems, or those that must obey complex symbol-based rules, such as buying stocks, filling out crossword puzzles, or choosing materials for a tool. Human choice is most commonly a culturally determined process in which many basic cognitive operations play a part, and the mechanisms of such decisions must be regarded as hybrid systems in which both brain and culture play a role. When the individual "makes" a choice, that choice has usually been made within a wider framework of distributed cognition, and, in many instances, it is fair to ask whether the decision was really made by the distributed cognitive cultural system itself, with the individual reduced to a subsidiary role.

This is not a simple issue to resolve. The individual brain and mind contain within them a great deal of history and structure that can be brought to bear on how decisions are made in specific cases. The differentiation of brain regions engaged in a specific kind of decision may be determined as much by this history as by innate brain growth patterns, both because of the epigenetic impact of culture on cognitive architecture, and because the actual task imposed on the brain by a given decision can be changed by redistributive effects that occur within the networks of culture.

In this way, distributed systems are able to change where in the system each component that influences a certain decision is located. This applies to such things as the locus of memory storage for a specific item, the locus of a specialized cognitive operation, and the locus of the choice mechanism itself. When all these components are located in a particular individual brain, decision making is one thing; when they are distributed across the brains of various people and information media, it is quite another, even if the final "decision" is made by one person. For this reason, and because decision making is so wide in its application, we should not predict the existence of a specialized brain region or subsystem that is devoted to resolving decisions in any general sense of the term. Rather, decisions can be made in a variety of ways that involve different anatomical subsystems of the brain.

Choice may be made automatically, without conscious engagement, in highly overpracticed routine situations. But, even in such situations, when necessary, an actor can intervene consciously at any level in the cognitive hierarchy of control. Thus, a pianist might play a well-rehearsed piece automatically but might also choose to consciously modify his playing online at the level of finger position, heaviness of touch, phrasing, formal interpretation, or emotion, depending on the feedback he gets from listening to his own performance. The important conclusion is that the possibility of conscious intervention in every choice remains viable, even after the automatization of a particular skill.

Certain kinds of choices must be made consciously—especially those involving novelty or learning. But, for the larger distributed systems of culture to operate smoothly, the rule seems to be: the less conscious engagement, the better. Conscious intervention is needed for acquisition and feedback control, but most individual cognitive operations should be made as automatic as possible in such systems. This constitutes a kind of industrialized cultural-cognitive network coordination, with concomitant efficiencies and increased collective power.

The superior powers of spontaneous intuition are often cited in studies of decision making (cf. Engel & Singer, 2008). Such intuitions are sometimes considered unconscious. But this is a rigged argument. In most examples, these cases of superior intuitions are clearly the outputs of deeply enculturated routines that have been learned by means of extensive and repetitive conscious rehearsal and refinement. This applies in principle to examples such as face recognition (an acquired, highly culture-specific skill), chess playing, various kinds of social judgment, business decisions, the composition of poetry, and so on. In each case, the task hierarchy is typically acquired in the context of a distributed cognitive-cultural system, with extensive pedagogy and training, and highly conscious, or deliberate, practice and rehearsal (sometimes in the form of imaginative play), to the point of automatizing the response, which affords the temporary release of conscious monitoring. In short, one of the main objectives of human conscious supervision is to make itself less important in future performances of the same activity.

However, as in all things, there is an associated caveat: While there are obvious advantages to abandoning personal conscious control in overlearned performances, under some circumstances (especially in novel situations), there are equally obvious dangers. The conscious individual is still the ultimate arbiter of choice, and the systems that remember networks of cultural practice and transmit their functions to new generations cannot be constructed without a great deal of conscious deliberation. This is where freedom gains an especially meaningful degree of traction. The advantages of a loss of conscious control in many short-term actions should not overshadow the oversight role of consciousness in creating and maintaining the elaborate cognitive apparatus of culture, keeping it on course, and preventing it from becoming unstable. Call it distributed metacognition. Whatever the label, it is at the very heart of human freedom to choose.

## DISCUSSION WITH

*Might there be a*
*possibility space of*
The main feature of
the invasion of the
into personal experience.
is to expose them
manner. Interactive
train up attention and
in this society. In addition,
skills that will be helpful
so characteristic of this

There are probably
experience at such a
mental skills that help
experiences afforded by
does not interfere with
constructive.

*Would a robot that*
*candidate for consciousness*
Yes. In principle, it is
However, the criteria for
tivity, curiosity, and
tionary perspective. Other
developing them.

Rather than contempla
to consider the kinds of
anywhere close to meeting
able of extraordinary
curious or autonomous,
meaning found in human

*How are humans different*
In many ways, which I
important difference that
matized attentional manage
taining. The neural basis for
with the selective evolution
which is around 10 times

Humans and other
cognitive capabilities,

## DISCUSSION WITH MERLIN DONALD

*Might there be a way to aid children in learning how best to navigate the possibility space of modern conscious experience?*
The main feature of modern conscious life that stands out from the past is the invasion of the highly engineered products of the electronic world into personal experience. One possible way to help children deal with this is to expose them early to this environment in a benign and protected manner. Interactive computer games and even use of the Internet can train up attention and memory management skills that are very important in this society. In addition, this kind of experience develops metacognitive skills that will be helpful in dealing with the rapid turnover of information so characteristic of this age.

There are probably trade-offs in placing emphasis on this type of experience at such a young age, but it does seem to cultivate the kind of mental skills that help meet the challenges presented by the "engineered" experiences afforded by the new electronic environment. So long as this does not interfere with their social development, it might be seen as constructive.

*Would a robot that demonstrated sensitivity, curiosity, and autonomy be a candidate for consciousness?*
Yes. In principle, it is possible that a robot might meet these criteria. However, the criteria for consciousness referred to in this paper, (sensitivity, curiosity, and autonomy) were developed mainly from an evolutionary perspective. Other species, and not robots, were the context for developing them.

Rather than contemplating imaginary robots, it might be more useful to consider the kinds of robots that currently exist. These do not come anywhere close to meeting these criteria. Some existing robots are capable of extraordinary sensory and perceptual sensitivity, but none are curious or autonomous, and none has the sensitivity to nuances of meaning found in human beings.

*How are humans different from other species?*
In many ways, which I have discussed at length in my two books. One important difference that is often overlooked is the high speed of automatized attentional management that human beings are capable of sustaining. The neural basis for this capacity probably has some connection with the selective evolution of the frontal-cerebellar tract in humans, which is around 10 times larger in us than in chimpanzees.

Humans and other species are qualitatively different in their surface cognitive capabilities, such as language and other skills. But, more fundamentally, humans are fantastically fast and nimble at automatized attentional management of all kinds. It is the capacity for this kind of attention that stands out as a cognitive precondition for the complex behaviors that define us. Piecing together a conversation, or playing a game of baseball or soccer, is a feat that requires a unique high-speed, multchannel, automatized attentional system. This is why even chimpanzees cannot match the skills of children at such things.

*Are attention and consciousness the same thing?*

No. Complex cognitive actions involve an attentional hierarchy. Some of the attentional acts in the hierarchy are conscious, and some are not. One example that illustrates the difference between consciousness and attention is reading. When children learn to read, much of their early training is concerned with where to look, and what to notice on the page. This requires conscious control of visual attention. However, once children have learned to read, their eye movements no longer need to be consciously controlled, at least under normal circumstances. The focus is on meaning.

Nevertheless, even in skilled readers, visual attention remains essential to the act of reading. There is a complex and specific sequence of eye movements that enable the brain to integrate what otherwise would be confusing images into a meaningful text. The point is that in a literate person those eye movements become part of an automatized attentional hierarchy that controls the act of reading. The hierarchy of attention is thus not entirely "in" consciousness. The eye movements still constitute an attentional act, and yet they are no longer in the focus of awareness (of course, if necessary, they can return there). The subject is normally "aware" only of the top end of the hierarchy of attention. Attentional acts at lower levels in the reading hierarchy are not fully conscious, but nevertheless operate effectively at some less conscious level of the mind.

## REFERENCES


Baars, B. J. (1988). *A cognitive theory of consciousness*. New York: Guilford Press.

Bargh, J., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin, 126*, 925–945.

Brooker, B. H., & Donald, M. (1980). The search for scalp-recordable speech potentials. In H. Kornhuber and L. Deecke (Eds.), *Progress in brain research, Vol. 54, Motivation, motor and sensory processes of the brain: Electrical potentials, behaviour and clinical use* (pp. 782–789). Amsterdam, The Netherlands: Elsevier, North Holland.

Dennett, D. (2003). *Freedom evolves*. New York: Viking.

Donald, M., Inglis, J., Sproule, M., Young, M., & Monga, T. (1980). The averaged movement potential in stroke hemiplegia: Comparison of intact and pathological

hemispheres. In
*Vol. 54, Motiv*
*behaviour and cli*
North Holland.

Donald, M. (1991). O
*and cognition*. Ca

Donald, M. (2001). *A*
W. W. Norton.

Donald, M. (2008).
distributed cogni

Donald, M. (in press).
(Eds.), *Becoming h*
UK: McDonald In

Edelman, G., & Tono
*imagination*. New

Engel, C., and Singer,
*and institutional a*

Hebb, D. O. (1963). Th
*Psychologist, 18*, 1

Kornhuber, H., &
Wilkurbewegunge
Beritschaftspotent

Libet, B. (1985). Unc
voluntary action. B

Libet, B. (1993). *Neuro*

Nelson, K. (1993). Ps
*Psychological Scienc*

Nelson, K. (1996). *Lan*
*mind*. New York: C

Newman, J., & Baars,
consciousness. *Con*

Schaller, S. (1991). *A m*

Walter, W. G., Coope
(1964). Continge
association and exp

Wegner, D. (2002) *The*

hemispheres. In H. Kornhuber and L. Deecke (Eds.), *Progress in brain research, Vol. 54, Motivation, motor and sensory processes of the brain: Electrical potentials, behaviour and clinical use* (pp. 487–490). Amsterdam, The Netherlands: Elsevier, North Holland.

Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Cambridge, MA: Harvard University Press.

Donald, M. (2001). *A mind so rare: The evolution of human consciousness*. New York: W. W. Norton.

Donald, M. (2008). The slow process: A hypothetical cognitive adaptation for distributed cognitive networks. *Journal of Physiology (Paris), 101*, 214–222.

Donald, M. (in press). The exographic revolution. In C. Renfrew and L. Malafouris (Eds.), *Becoming human - Innovation in material and spiritual cultures*. Cambridge, UK: McDonald Institute, University of Cambridge.

Edelman, G., & Tononi, G. (2000). *A Universe of Consciousness: How matter becomes imagination*. New York: Basic Books.

Engel, C., and Singer, W. (2008). *Better than conscious? Implications for performance and institutional analysis*. Cambridge, MA: The MIT Press.

Hebb, D. O. (1963). The semi-autonomous process; its nature and nurture. *American Psychologist, 18*, 16–27.

Kornhuber, H., & Deecke, L. (1965). Hirnpotentialandeerungen bei Wilkurbewegungen und passiv Bewegungen des Menschen. Beritschaftspotential und reafferente Potentiale. *Pflügers Archiv, 284*, 1–17.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences, 8*, 529–566.

Libet, B. (1993). *Neurophysiology of consciousness*. Boston: Birkhauser.

Nelson, K. (1993). Psychological and social origins of autobiographical memory. *Psychological Science*, 4, 7–14.

Nelson, K. (1996). *Language in cognitive development: The emergence of the mediated mind*. New York: Cambridge University Press.

Newman, J., & Baars, B. J. (1993). A neural attentional model for access to consciousness. *Concepts in Neuroscience, 4*, 255–290.

Schaller, S. (1991). *A man without words*. New York: Summit Books.

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain. *Nature, 203*, 380–384.

Wegner, D. (2002) *The illusion of conscious will*. Cambridge, MA: MIT Press.