

Studying Consciousness Using Direct Recording from Single Neurons in the Human Brain

Moran Cerf and Michael Mackay

Abstract Consciousness represents a unity of experience derived from limitless possibilities and fierce competition for dominance within our brain. This unity reflects the compromise of an internal representation of the world that shapes our perception of the environment and guides the choices we make in interacting with it. Here we investigate the relationship between the external environment, our internal representation of it, and volition, using direct recording from single neurons in the brains of 12 patients with pharmacologically intractable epilepsy. In two novel experiments we pit the world within against the world outside, and separate the decision making processes within the brain from the sensation of volition. We show that subjects are capable of overriding external sensory input with internal imagery, and can directly control the firing rate of individual neurons in the medial temporal lobe. Additionally, we show that decision can be predicted before the conscious perception of volition, but that feeding this decision forward and bypassing volition leads to a gradual loss of predictive ability as the neurons alter their activity in response. We propose a working definition for consciousness based on these results, suggesting that consciousness involves our memory and experience of events and ties these together in a working narrative.

1 Introduction

Studies in psychology separate two entities competing in our brain for dominance. One is the “experiencing I,” characterized as “living in the present,” intuitive, and effortless. This entity, commonly termed “system 1,” is automatic, opaque, skilled

M. Cerf (✉)

Computation and Neural Systems, California Institute of Technology, Pasadena, CA 91125, USA
and

Department of Neurosurgery, University of California Los Angeles, Los Angeles, CA 90095, USA
and

School of Business, New York University, New York, NY 10012, USA

e-mail: moran@morancerf.com

and associative. The other entity, “system 2,” is the “remembering I,” which is deductive, slow, self-aware and top-down controlled (Sloman 1996). The latter is the one we commonly interact with consciously. When we are about to choose a vacation, for instance, we put the two to competition: the experiencing system wants to go to a new place and generate more experiences to live by, while the remembering one wants to choose the best memory from a list of existing memories – a memory that was already proven to be positive.

These systems reflect a fundamental neurological separation of the outside, sensory world from its internal representation in our brain (Cerf et al. 2010). The ability to separate that which is internally generated from that which is externally generated is arguably what underpins our identity (Frith 1987). Neuroscience often strives for a view of the world where a representation of the environment exists directly within our brain – a “what you see is what you get” type of phenomenon. The outside world is then represented in our brain as is. But in life we know that we can shape the information in our mind in multiple ways: information from the environment can be stored in our brain in a less informative set of values than the actual experience (Redelmeier and Kahneman 1996), information retrieved from our memory can be reconsolidated with changes (Schiller et al. 2010), and information can be manipulated by our senses (Cerf et al. 2010). Thus we find a blurring of the lines between the internal representation of the world and its external manifestation. This separation is conceptual but also exists neurologically in the form of different regions within the brain.

While our brain can be thought of as a coalition of various sub-parts, we normally think of our personality as a coherent entity. That said, we know that these sub-parts occasionally vie for dominance within the brain (Kreiman et al. 2002). Studies such as the ones conducted with split-brain patients – patients whose right and left hemispheres were separated in surgery for clinical reasons – explicitly demonstrate this competition because the two hemispheres can be given conflicting tasks. The two halves of the brain, when separated, act as two individual entities, each controlling their respective half of the body and responding to their half of the sensory world (Gazzaniga 2005). Such phenomena could happen in all our brains with binocular rivalry tasks (Leopold and Logothetis 1996), where different images which are presented to the two eyes yield conflicting percepts. However, in normal subjects this competition is resolved moment to moment to give a single conscious perception, despite its fluctuant nature (Levelt 1965).

The actions we project to the world are the sum of internal ruminations and deliberations, which are commonly aggregated into a single decision. Once a decision is made, there then occurs one of the most debated phenomena in philosophy of mind: volition. Though impossible to discern in its metaphysical guise, modern neuroscience has opened the fundamental question of “free will” up to scientific investigation by looking for the signature of decision and motor initiation in the brain. To do so a range of techniques has been used: from direct electrical stimulation to cutting edge neuroimaging and animal neurophysiology. Evidence of a volitional signal emerged with the recording of the *bereitschaftspotential* (Kornhuber and Deecke 1965; Libet et al. 1983; Libet 1999), which was a rising

level of activity in motor cortex preceding voluntary muscle movement. Scalp EEG and MEG experiments with humans then revealed changes in neural activity that preceded awareness of volitional state by hundreds of milliseconds (Haggard 2008).

We now know that there are multiple stages to motor initiation: the rise of a “readiness potential,” the conscious urge to act, and the action itself (Fried et al. 2011). While the initial change in neural activity can precede the conscious urge by as much as seconds, the urge and the action are typically separated by 100–200 ms.

These separate stages allowed us to probe two intriguing aspects of volition: firstly, can a volitional signal be decoded on a single neuron level, and secondly, what happens when the brain is made aware of its decision before the conscious observer is? Furthermore, we were interested in seeing what changes could be incurred in the brain following direct tampering with its ability to control the flow of information.

To address these questions, we designed two experiments that offered human subjects the ability to alter the activity of neurons in their brains while getting direct feedback on this activity. Here we explore how and where the interaction between our external and internal worlds occurs in the brain by directly pitting the world within and the world outside against each other. In doing so, we directly test the ability of subjects to alter the dominance of brain regions and show evidence for the ability of humans to shape the way information is perceived in their brains.

2 Experiments

The subjects in these experiments were 12 patients with pharmacologically intractable epilepsy who were undergoing brain surgery to localize the seizure focus for possible surgical resection (Fried et al. 1997). During this procedure, the patients were implanted with intracranial electrodes accessing regions in the medial temporal lobe (hippocampus, amygdala, entorhinal cortex and parahippocampal cortex). After surgery, subjects participated in a series of experiments. In the first experiment, they were instructed to play a game in which they controlled the display of two superimposed images via the firing activity of four medial temporal lobe (MTL) neurons in the brain (Fig. 1a). In a prior screening session, we had identified four different neurons that responded selectively to four different images by increasing their firing rate when the corresponding image was presented. Each trial started with a two-second display of one of these four images (the target). Subjects next saw an overlaid hybrid image consisting of the target and one of the three remaining images (the distractor) and were told to enhance the target (“fade in”) by focusing their thoughts on it. The initial visibility of both images was 50% and was adjusted every 100 ms by feeding the firing rates of four MTL neurons into a real-time decoder that could change the visibility ratios until either the target was fully visible (“success”), the distractor was fully visible (“failure”), or until 10 s had passed (“timeout”).

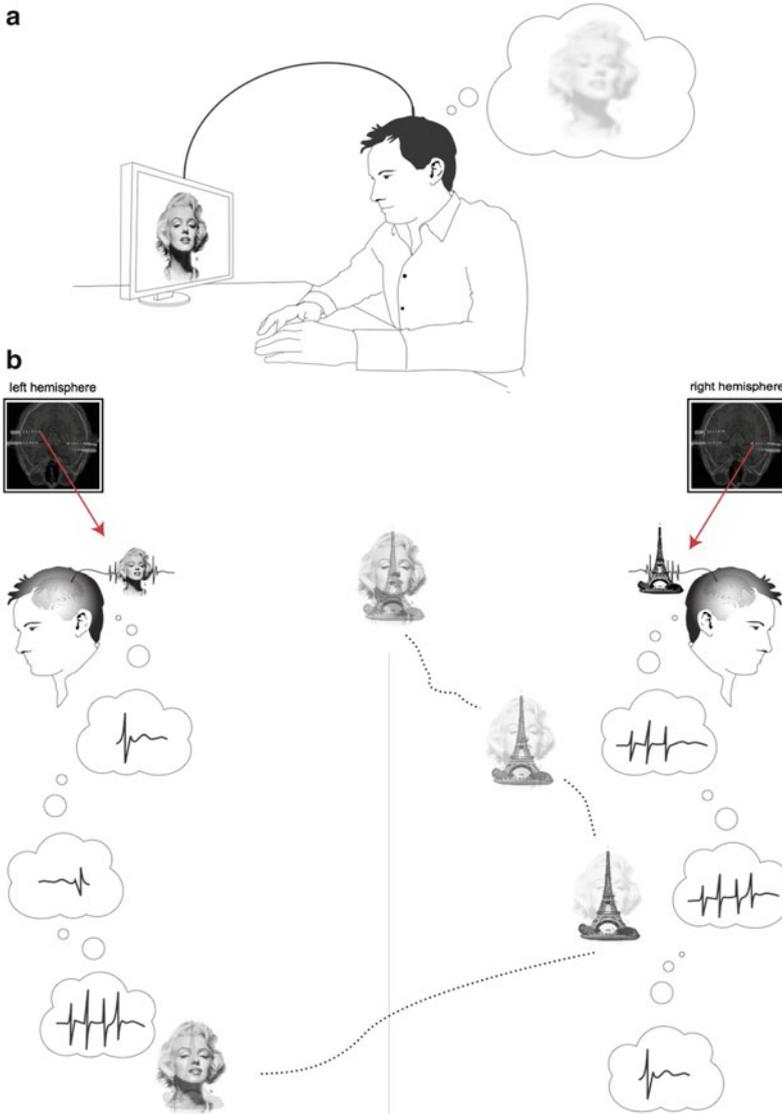


Fig. 1 (a) Recording from intracranial electrodes, neurons are identified that respond to a specific concept. In this instance a cell responsive to the image of Marilyn Monroe was found. This cell increases its firing rate to the image or thought of Monroe. (b) This cell is then pitted against one found to represent the Eiffel tower. The two images are superimposed and the subject is asked to bring the image of Monroe to maximum visibility. The visibility of the image is controlled by real-time decoding of the activity of each neuron relative to the other neuron and its own baseline. In this example, we show a case where the subject initially begins to fail the experiment – the firing of the Eiffel neuron increases and the visibility of the tower increases, creating negative feedback. However, the subject is able to exert control and, by concentrating on the internal thought of Monroe, is able to override this sensory input and increase the firing rate of the Monroe neuron and decrease that of the Eiffel neuron, bringing the image of Monroe to visibility. The scans show the location of the respective electrodes within the brain

The subjects manipulated the visibility of the hybrid images by any cognitive strategy of their choosing. Six of 12 subjects reported in a follow-up interview that they focused on the concept represented by the target picture or closely allied associations. As they performed the task, we showed that subjects could control their own brains as though they were controlled by an external device (Cerf et al. 2010).

Occasionally the subject nearly failed in the task (Fig. 1b). While they were told to focus their thoughts on a particular concept (say, Marilyn Monroe), which was pitted against a different concept (say, the Eiffel tower), the distractor gradually became more dominant on the screen, nearly leading to a failure by becoming fully visible. However, just when the subject was about to fail the trial, he/she was able occasionally to summon the thought of the target concept into his/her mind, making the neuron or coalition of neurons corresponding to Marilyn Monroe engage in activity and shift the hybrid image back to the target. This finding implies that, while visual feedback showing an image of the Eiffel tower was penetrating the brain, the subject was able to override this information with the mental imagery of Monroe. Thus, this experiment has a deeper and more profound result: imagery can override vision at the level of the MTL, or as we choose to put it, idealism trumps realism. The world in the subject's brain can override the outside world.

Practically, this is equivalent to seeing a cup of tea with your eyes, but choosing, in Kantian terminology, to wear different glasses, such that the image of the cup is turned inside our mind to a picture of a flower. Our "flower" neuron would fire and we would actually "see" a flower, although the photons hitting our retina would reflect a cup of tea. The same analogy can be used for any stimuli: hearing a word differently than what it was when a person swore at you; choosing to ignore pain; or choosing to not let any external stimulation penetrate when you focus your mind within. We can synthesize an internal world to override the one given to us by the senses.

If this merging of the external and internal worlds is constantly happening, then one can ask what our conscious experience actually represents. From a purely sensory basis, we would want this information to be the most accurate representation of reality that our sensory systems are capable of generating. However, we experience, learn, and create an internal representation of the world, describing how we expect it to behave and how we interact with it. Thus the phenomenal reality we perceive is a blend of these two worlds – the objective, external world, shaped by our subjective, internal world – to create what is effectively a learned representation of reality.

In this study, we show evidence of humans' ability to alter the way information flows in the brain, using information from the visual system. Practically, this can be extrapolated to any procedure of experiencing the outside world. For any means of altering perceptual information, it is likely that our "remembering I" performs these kinds of modifications to our sensory input constantly, thus underlying our well-described predisposition to perceive the world we expect to perceive (Merleau-Ponty 1996).

An important note to address in this experiment is the question of what happens at the neuronal level. To answer that question, we measured the change in activity of neurons in the brain pitted together in this competition. A success in each trial in which two neurons are pitted against each other can be the result of three possibilities: the neuron selective for the target (target neuron) increasing its firing above its baseline, the distractor neuron in this competition decreasing its firing below its baseline, or the two neurons working together, with target rising and distractor declining. Testing this point directly shows that the third case is the most common. In over 80% of the trials, two neurons, often separated in two hemispheres and distant from each other, ended up working together to perform the task, with one neuron significantly increasing its firing rate while the other decreasing its own. This sort of entanglement between two neurons is profoundly surprising. The two neurons, along with the subject, had no knowledge prior to the experiment which two images would end up being placed in competition. The images were selected based solely on the morning's screening. Moreover, each trial pitted together a different pair of images not known to the subject in advance. However, as soon as the trial began and the subject realized which two images were competing, they immediately controlled one neuron to go up while the other went down, channelling information far into two different locations rapidly.

A question that comes to mind is what coordinates this activity such that both neurons simultaneously alter their activity to work together to perform the task so rapidly and profoundly? We know from testing sets of neurons all firing for the same concept that the effect is not purely working on an individual neuron but on a small coalition of neurons all responding to the same concept (Cerf et al. 2010). We also know that this effect is very sparse (Waydo et al. 2006). Only the Monroe and Eiffel neurons will alter their behavior and not a larger region of the brain (Cerf et al. 2010). Our results show that neurons adjacent to the Monroe/Eiffel neurons will not change their firing patterns in this trial; only the two relevant clusters of neurons will. More specifically, in a following trial, we could pit Monroe with a third image, say, an image of Johnny Cash. In this trial, Monroe neurons would go up while Cash neurons go down, but the Eiffel tower neurons from the trial before would remain untouched. If in a following trial Cash and the Eiffel tower would compete, the Monroe neurons will remain still while the Eiffel/Cash neurons go up/down or down/up, based on the target of the trial.

These changes happen rapidly and accurately, and the specificity suggests that this effect is not simply the result of lateral inhibition. These results fit with those seen in neurophysiological studies of attention in primates (Reynolds et al. 1999), in that the change in response is seen after a delay and it increases if the neuron represents the attended image but decreases if that neuron's target is not attended to. This finding suggests that the coordinators of such changes in response may be attentive processes, which makes sense in the context of the subject's reported cognitive strategies.

Finally, it is important to note that the characterization, properties and behavior of these neurons seem very similar to those of place cells in rodents. If one assumes that thoughts in "concepts space" are as relevant to humans as locations in "place

fields,” and that humans practically navigate their lives within such a concept space as rodents navigate in a maze, then one can draw this analogy between the properties of the two and suggest that the MTL is the seat of the dominant entity that we use to navigate by in our minds. Humans travel inside their thoughts as rodents navigate the world in a maze. Our two-images competitions show strong analogy to the results of studies showing experiments in which rodents were given a two-way maze to choose from and the typical firings of two single place cells were pitted against each other in individual trials (O’Keefe and Nadel 1978; Wilson and McNaughton 1993; Eichenbaum 1996; Moser et al. 2008).

As the first experiment demonstrated the subjects’ ability to increase or decrease the activity of single neurons in their brains voluntarily using their sheer will, we wanted additionally to consider what changes could occur in the same neurons without conscious control but based on feedback from the environment. To do so, we designed an experiment that separated the subjects’ desire to act from the neuronal mechanisms that precipitated the feeling of volition. Following a prior experiment, which showed that it is possible to decode the activity of single neurons in subjects’ brains prior to their volition to act (Fried et al. 2011), we designed an additional on-line experiment that interacted with these neurons prior to the subjects’ conscious knowledge of the act.

The subjects were faced with two red buttons and were asked to press either the left or right one with their hands, in their own time. Each trial started with the subject having to wait for three seconds until a “get ready” sign disappeared. Once the signal disappeared, the subjects had as much time as they wanted to feel the urge to press a button before choosing which one to press, and as they did so the button was “lit” (turned on). The subjects then waited 3 s for the light to turn off before another trial began. Occasionally, the button was turned on without the subject pressing it. The subjects were told that this sometimes happens and were asked to wait for the button to be turned off again and for the following trial to start. Subjects were not encouraged to press the button quickly or slowly but were specifically told to wait until they felt the urge as need be. They were only instructed not to press the button when the light is on.

The first 50 trials were used to train a decoder on identifying sets of neurons that fired prior to the subject’s urge to move. We identified neurons that fired between 1,500 and 200 ms prior to the subject’s movement and between 1,000 and 100 ms prior to the subject’s reported “will” to initiate a movement. Once the computer identified these neurons, or a classifier passed the threshold of 75% accurate predictions of the following button press at least 300 ms in advance, we began the active part of the experiment. While the subject still performed the same task, we altered the mechanism that turned the button on. Contrary to the original trials, in which the choice to turn the button on was either based on a press of the button or random timing, in the second phase we turned the button on either randomly (30% of the trials) or based on a decoding of the subject’s will to press a button (Fig. 2).

As we decoded the subject’s will to press a specific button prior to their conscious perception of this will, we generated two more conditions whereby we turned the predicted button on at different times. Either we intentionally waited

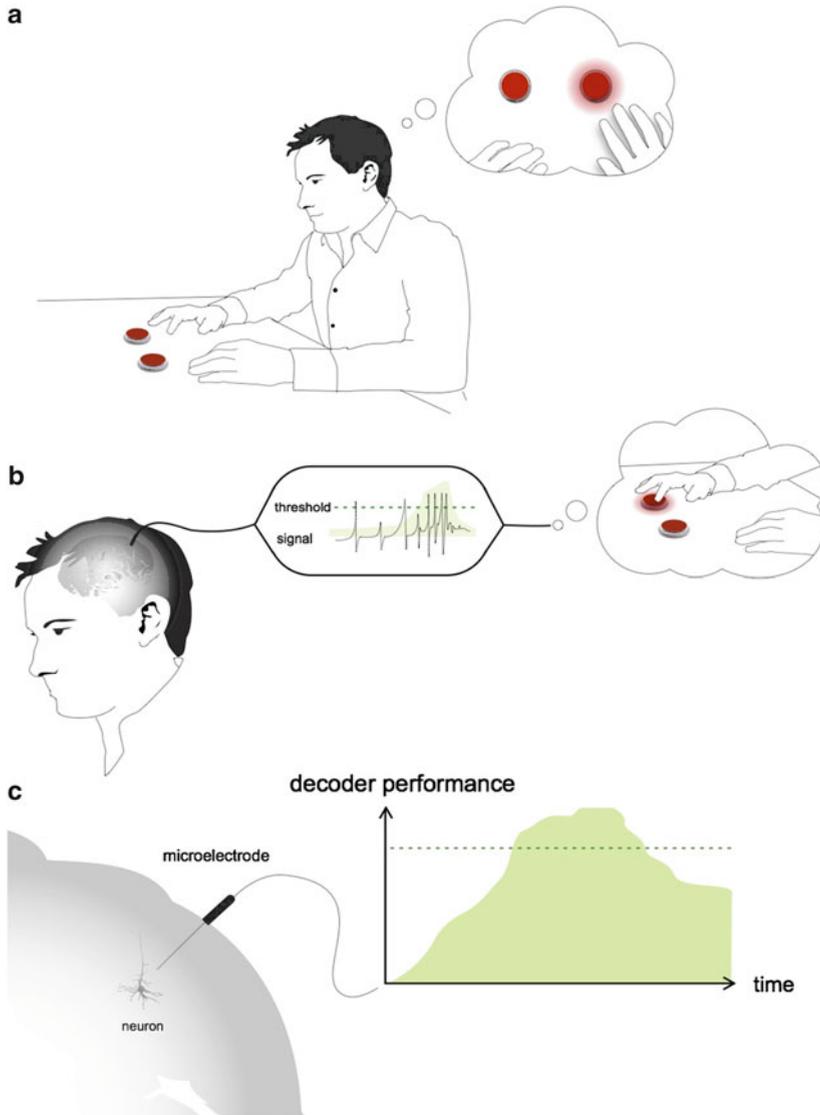


Fig. 2 (a) The subject is faced with two buttons. At the start of a trial, he/she is allowed to choose at will which button to press when he/she feels the urge to do so. On pressing the button, it lights up. In some trials a button lights up prior to the subject pushing it, based on the decoding of the urge to press the button. (b) Selected neurons show an increasing rate of activity until a threshold is reached at which the subject feels the urge to move, but before motor initiation. (c) For the initial trials, this activity of neurons predicting volition was decoded until capable of accurately decoding the subjects' choice prior to the conscious urge to move. Once the threshold of 75% accuracy was reached, we inserted a condition in which we used this information to light the predicted button up prior to the subject feeling of urge to press it. After a number of trials, the success of the decoder started to reduce as the neuron altered its activity, presumably in response to being decoded

500 ms (the presumed time it takes the subject to become aware of his/her will to press the button) and then turned the button on, or immediately turned the button on, such that the subject was unaware of his/her brain's presumed will to press the button. The third condition created a unique situation in which brain activity gave rise to an action in the environment without the subject's conscious awareness that he/she had triggered it.

From the subject's point of view, the "random" trials and the predicted activation trials were identical. In both cases, the button was turned on without the subject feeling that he/she had any effect on it. The difference, though, is in the "neuronal" view of the experiment. In this case, the neurons (to anthropomorphize them) did know that they were active and were signalling their "will" to press the button. The feedback from the environment confirmed to these neurons that what they wanted was done; however, the internal feedback said that the result was not triggered by the subject.

Our results show that, on the single neuron level, looking at the decoding ability of an individual neuron that is identified as a key module in giving rise to volition, the neurons reached a plateau of maximal decoding within the first trials and were being decoded in real time during the following trials. However, we saw a gradual decrease in the performance of these neurons as the experiment progressed. Simply put, the neuron learned that it was being decoded and that the decoding interfered with the task at hand and, within a few trials, learned to not allow for the decoding, all without the subjects' knowledge. The subjects showed no conscious knowledge of the shift in decoding or in the ability to affect the trials that were based on the neuron's activity, but we saw that the decoder's performance was dropping based on the success of the decoder to interfere with the neuronal activity. Practically, the subjects just claimed to have no desire to press the button for a few trials (the trials within which the button was pressed because of their unconscious will), after which they regained the knowledge of their will and again felt the desire to act.

It may simply be that, while these neurons initially signal that they want a specific choice to be made, if there is repeatedly a lack of feedback that their "will" has been effected by the subject then they may reduce in activity.

3 Discussion

In the first experiment, we demonstrated the subjects' ability to up- and down-regulate the activity of neurons voluntarily using their thoughts. In the second experiment, we showed neurons changing their behavior based on a disparity in feedback from the internal and external worlds. Both experiments demonstrate direct access to the activity of single neurons in the brain without changing the entire sequence of activity of neighboring neurons. Moreover, both experiments show evidence of a change in network activity based on the task, altered either by attention or by the effect of neuronal response on the environment. However, while the first experiment happens with conscious and direct control arriving from the subject, the second one bypasses the subject's awareness. This raises a question:

when neuronal alteration of behavior occurs in the brain due to feedback, who is getting the feedback and exerting the change?

It might appear as if there are two people involved in these experiments. One is the subject's mind, instructing the brain to think of a concept, while the other is the one that is acting out the mind's desire, namely, the nerve cells in the MTL that up- and down-regulate their activity accordingly. A fundamental peculiarity of the mind-body nexus is apparent here: we do not think in terms of neuronal firing rates. Our thoughts, desires and feelings are metaphysical entities. We have no knowledge of the mechanisms that interact inside our brains when we act. Yet the thought of a concept such as Marilyn Monroe translates into a particular pattern of neuronal activity when we elicit her image in our minds (Kreiman et al. 2000). Events in the phenomenal mind find their parallel in the material brain. We have to wonder, however, if it is our conscious desire to "see" Monroe that drives the change in neuronal firing or the change in neuronal firing that generates our desire. Who is the "puppet" and who is the "puppeteer?"

Practically, it is the brain that fulfils both roles; the same brain is learning about the task and performing it. The "puppet" is also the "puppeteer" and most likely represents a circular route of causality. There is a singular level of neuronal coalitions that represent conscious perceptions that is subject to incoming sensory information and is then appraised relative to our internal world and desires, and which then feeds back and modifies the activity of these same neurons. Incoming information then simply feeds into this eternal loop.

One can ask then, when brains communicate, either with the outside world or, for that matter, with themselves, through internal ruminants or via imagery, what gives us the illusion that one person is "listening" while the other is "performing?" Why, when speaking to someone, despite both parties hearing the speech simultaneously, do we somehow feel that these thoughts are owned by us more than they are owned by others (Stephens et al. 2010)?

It is worth noting that the brain separates the internal world from the external all the time. In doing so, it attaches a profound sense of ownership to conscious sensations that it determines as internally generated, such as our internal voice or the sensation of moving one's body. Indeed, one of the most devastating psychiatric disorders, schizophrenia, can be explained in terms of the brain's inability to correctly separate out these two streams of sensory information (Frith 1992). Thus the internal voice is misattributed as an external sensation. At the same time, a fundamental loss of sense of self may underlie thought interference, in which a person believes his/her thoughts are open to being read, or withdrawn, or inserted by another party; the person loses the profound sense of ownership we have of our own thoughts (Nelson et al. 2009).

In this context, the experiment in volition provides a remarkable window on the processes that underlie this attribution of ownership to sensory information about external events. We have a decoder that can accurately predict a decision before the conscious perception of will, but also before a motor command is generated. We also know that the subject feels no ownership of the button being lit up if it does so before they become aware of the will to press it (roughly 100–500 ms after the

neuronal activity the decoder reads). Thus we have bisected the process at a point where what has yet to happen is crucial to applying a sense of ownership to the button lighting up, but parts of the brain have already prepared and effectively selected the decision to come. This finding allows us to look at how the brain responds to such a contradiction but also potentially allows us, in the future, to dissect out what contributes to volition and the sense of ownership.

This relationship between the internal world and the external world is important in determining the content of consciousness. With that in mind, one might ask if we can access a person's consciousness without him/her being involved whatsoever. In effect, the advertising industry does this all the time: making us buy the soup that they want without noticing that we wanted it (we do not know any more if we like the taste of Coke; we just know it as something we like). In the same way, we were able to implant in our subjects the occasional firing of a neuron either with the aid of their conscious volition or with the aid of the prior mechanisms that gave rise to that sense of volition, even without an active feeling of agency; without generating the thought in their mind.

The reason for the discrepancy between the conscious feeling of agency in the first experiment and the unconscious one in the second could be attributed to memory. Memories are an essential second component of consciousness. Memories tie together the set of stories we have into a narrative. That is because the thought requires an additional component on top of the activity of these neurons: it requires a narrative to be built from it. The sensory world we are immersed in is translated into neuronal representations, and these representations are dependent on past experiences, our internal world and its narrative. The essence of consciousness is the material that selects the one percept from the many that rises to our awareness. To that effect, our experiment is a demonstration of humans' ability to control what is accessible to their minds. Our subjects could practically choose to not see things that were visible to their eyes and see things that were not, purely based on thoughts. In doing so, they could actively select which patterns to encode in their memory and which not to.

This is almost a Buddhist point of view: the world in front of our eyes is just a suggestion whereas the world inside our brain is ours to create. Put differently, while some aspects of the world seem to occur to us, much of it is ours to construct. The person cursing you on the street because you cut him off driving is sending an auditory signal that will most certainly be processed by your ears and triggering activity in your auditory system, but whether you let it alter your consciousness or not is up to you and the narrative you are building. When we walk in the streets and see three little kids walking towards us, the photons from the environment are hitting our eyes simultaneously. But only a complex circuit in one's brain can lead to "noticing" one's son on the left first and not the other two children to his right. Presumably, for the same visual input a different parent will notice the girl in the center first and will give rise to a different neuronal representation despite an initial visual response that is nearly identical.

Additionally, these two experiments touch on the question of what parts of our brain we are capable of exerting conscious control over. While subjects in our

experiment exerted control over brain regions such as the hippocampus and parahippocampal cortex, the same effect was not seen in great magnitude in other regions such as the amygdala or entorhinal cortex (Cerf et al. 2010). The latter showed an average performance level of 20% compared to nearly 70% in the hippocampus and 69% in the parahippocampal cortex. The ability or lack of ability to control our brains in regions such as the amygdala remains open for further investigation. The reasons behind the lack of control in some regions are not clear yet. However, one can propose that this lack of control is tied to our general conception of hardship in controlling our emotions. With the amygdala being implicated as a primary seat of emotions, and our experiments showing a decreased level of ability to consciously control the amygdala neurons voluntarily, one can propose that this is evidence of the lack of ability to regulate emotions in some brain regions directly. While adult humans are able to regulate their emotions to some extent (unlike young humans or most animals which – for better and for worst – seem to be in the moment at any given time) we do not seem to show the same level of control over those as we show over our memories and decisions. We can summon a thought to our mind rather easily, but cannot control emotions as easily. We cannot choose when to feel things and exactly how and what to feel. Emotions somehow seem to dawn on us. They capture us and we steer our boat through them. At the extreme end of the spectrum of control, we have those who suffer from bipolar affective disorder, whose emotions are inherently unstable, uncontrollable, and excessively powerful.

As the question we posed in the beginning addressed not only the experimental results from studies of single neurons but also their relevance to the study of consciousness, we want to end by suggesting a working definition of what consciousness is and why it is something that we can probe best using studies of with single neurons in humans. Consciousness is the narrative that we compose about our actions. It is the story we make from the series of occurrences and events in our brain, in the form of past memories, current feelings and thoughts, simulations of the future, and sets of influences from the environment, be it via the five senses, internal rumination, or random guided processes that we attribute to our own free will. The major purpose of consciousness, therefore, is to communicate this story, this narrative, to as many equal brains as possible. The way to immortalize ourselves is by creating multiple instances of our existence, as accurately as possible, in as many brains as possible.

The idea of neurons encoding us (like the Marilyn Monroe neurons) could be, in that same point of view we remarked on beforehand, the ultimate way to sustain our existence. Having children and creating a new generation is such a guaranteed imprint of ourselves in an additional brain. Interaction with other humans and living organisms and accurate communication are the methods by which we ensure that our imprints and our experiences, as they are imprinted in other brains, are as accurate as possible. It is how we evolve ourselves and make sure we keep improving the source of those imprints in such a way that will allow us to increase the possibilities of future encoding of ourselves in more brains. The smarter we

become and the better communicators we become, the more likely we are to generate more such instances and more accurate ones.

We walk our story, talk our story and even live it when we are asleep. Every occurrence, even the ones we are not aware of, or the ones we prefer to not have experienced, shapes this narrative and offers a reflection on our existence that could lead to improvement in our behavior. Consciousness, therefore, requires knowledge of the source of its perceptions and for those perceptions to be consistent in the context of its ongoing narrative; it wants to always have an answer to what is happening regarding itself. It is the reason we always come up with an explanation for situations and the reason we try to constantly regulate our emotions and find it hard to accept that they dawn on us. It is the reason we suffer whenever we feel that our free will is in jeopardy, and the reason we conduct experiments that teach us how to control our own neurons. It is a step towards understanding the mechanisms to get closer to our story and make sure we can back it up after our bodies cease to exist. It is the main reason humans are so baffled by consciousness and the reason they spend so much time and energy using their brain to reflect and ponder about itself, study its own behavior, write papers and draw conclusions about itself. It is a quest for the nature of oneself.

Acknowledgment We wish to thank the patients for their participation in these studies; Daniela Schiller, Liat Maoz and Udi Pladott for help with the manuscript, and Maria Moon for help with the illustrations; Sonia Le Corneic and Yves Christen from the IPSEN foundation for support of this work. This paper is based on work conducted with Profs. Itzhak Fried and Christof Koch (Cerf et al. 2010).

References

- Cerf M, Thiruvengadam N, Mormann F, Kraskov A, Quiroga RQ, Koch C, Fried I (2010) On-line, voluntary control of human temporal lobe neurons. *Nature* 467(7319):1104–1108
- Eichenbaum H (1996) Is the rodent hippocampus just for place? *Curr Opin Neurobiol* 6(2):187–195
- Fried I, MacDonald KA, Wilson CL (1997) Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron* 18(5):753–765
- Fried I, Mukamel R, Kreiman G (2011) Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron* 69(3):548–562
- Frith C (1987) The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychol Med* 17(03):631–648
- Frith C (1992) The cognitive neuropsychology of schizophrenia. Lawrence Erlbaum Associates, Hove, UK
- Gazzaniga MS (2005) Forty-five years of split-brain research and still going strong. *Nat Rev Neurosci* 6(8):653–659
- Haggard P (2008) Human volition: towards a neuroscience of will. *Nat Rev Neurosci* 9(12):934–946
- Kornhuber HH, Deecke L (1965) Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflügers Archiv Eur J Physiol* 284(1):1–17
- Kreiman G, Koch C, Fried I (2000) Imagery neurons in the human brain. *Nature* 408(6810):357–361

- Kreiman G, Fried I, Koch C (2002) Single-neuron correlates of subjective vision in the human medial temporal lobe. *Proc Natl Acad Sci USA* 99(12):8378
- Leopold DA, Logothetis NK (1996) Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* 379(6565):549–553
- Levelt WJM (1965) On binocular rivalry. Institute for Perception RCO-TNO, Soesterberg, Netherlands
- Libet B (1999) Do we have free will? *J Consciousness Stud* 6(8–9):47–57
- Libet B, Gleason CA, Wright EW, Pearl DK (1983) Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): the unconscious initiation of a freely voluntary act. *Brain* 106(3):623
- Merleau-Ponty M (1996) *Phenomenology of perception* (trad: C. Smith). Motilal Banarsidass, Delhi
- Moser EI, Kropff E, Moser MB (2008) Place cells, grid cells, and the brain's spatial representation system. *Neuroscience* 31(1):69
- Nelson B, Fornito A, Harrison BJ, Yücel M, Sass LA, Yung AR, Thompson A, Wood SJ, Pantelis C, McGorry PD (2009) A disturbed sense of self in the psychosis prodrome: linking phenomenology and neurobiology. *Neurosci Biobehav Rev* 33(6):807–817
- O'Keefe J, Nadel L (1978) *The hippocampus as a cognitive map*. Clarendon, Oxford
- Redelmeier DA, Kahneman D (1996) Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. *Pain* 66(1):3–8
- Reynolds JH, Chelazzi L et al (1999) Competitive mechanisms subserve attention in macaque areas V2 and V4. *J Neurosci* 19(5):1736
- Schiller D, Monfils MH, Raio CM, Johnson DC, Ledoux JE, Phelps EA (2010) Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature* 463(7277):49–53
- Sloman SA (1996) The empirical case for two systems of reasoning. *Psychol Bull* 119:1.3–22
- Stephens GJ, Silbert LJ et al (2010) Speaker-listener neural coupling underlies successful communication. *Proc Natl Acad Sci USA* 107(32):14425
- Waydo S, Kraskov A, Quiñero R, Fried I, Koch C (2006) Parse representation in the human medial temporal lobe. *J Neurosci* 26(40):10232
- Wilson MA, McNaughton BL (1993) Dynamics of the hippocampal ensemble code for space. *Science* 261(5124):1055