Are we studying consciousness yet?

Hakwan C. Lau\textsuperscript{1,2}

\textsuperscript{1} Wellcome Trust Centre of Neuroimaging, University College London
\textsuperscript{2} Department of Experimental Psychology, University of Oxford

\textit{(‘near-final’ draft)}

To appear as a chapter in a Frontiers of Consciousness, edited by Martin Davies and Larry Weiskrantz (OUP, 2007/8)

Correspondence:
Hakwan Lau
Wellcome Trust Centre of Neuroimaging
12 Queen Square, London
WC1N 3BG, United Kingdom
Tel: 44 20 7833 7486
Email: hakwan@gmail.com
1. Introduction

It has been over a decade and half since Christof Koch and the late Francis Crick first advocated the now popular NCC project (Crick and Koch, 1990), in which one tries to find the neural correlate of consciousness (NCC) for perceptual processes. Here we critically take stock of what have actually been learned from these studies. Many authors have questioned whether looking for the neural correlates would eventually lead to an explanatory theory of consciousness, while the proponents of NCC research maintain that focusing on correlates is a strategically sensible first step, given the complexity of the problem (Crick and Koch, 1998; Crick and Koch, 2003). My point here is not to argue whether studying the NCC is useful, but rather, to question whether we are really studying the NCC at all. I argue that in hoping to sidestep the difficult conceptual issues, we have sometimes also missed the phenomenon of perceptual consciousness itself.

2. Stimulus Confound vs. Performance Confound

In neuroimaging, one standard strategy for identifying the NCC is to compare a condition where the subjects can perceive a stimulus against a condition where the subjects cannot. One example is visual masking, in which a visual target such as a word is rendered invisible by presenting irrelevant visual patterns before and after. Masking is an interesting paradigm because when the visual targets are masked and thus not seen, they are nonetheless presented to the retina. The retinal information is only rendered 'unconscious' because of the masks presented before and/or after it. Functional magnetic resonance imaging (fMRI) has been used to compare the neural
activity when subjects were presented with masked (and thus invisible) words, against the neural activity when subjects were presented with unmasked (and thus visible) words (Dehaene et al., 2001). It was found that, even when the words were masked, there was still activity in the visual areas, although its intensity was reduced. This means that in both the masked and unmasked conditions, the stimuli (the words) were present and to a certain degree processed in the brain. This is a critical aspect of the rationale of the experiment: what we are looking at is a difference in consciousness, but not a difference in the absence or presence of the external stimuli.

Some other paradigms allow us to maintain the constancy of the stimuli in an even more elegant fashion. In fact, much of our effort in the NCC project is spent dealing with this 'stimulus confound'. In other words, we try to make sure that any difference in the neural activity is not merely due to a difference in the stimulus. To this end neuroimagers employ various psychophysical paradigms. For example, in change blindness (Simons and Rensink, 2005) subjects frequently fail to detect a change in the visual scene. Researchers have compared the neural activity when the subjects detect the change with the activity when the subjects fail to detect the change (misses) (Beck et al., 2001).

Another elegant paradigm for maintaining stimulus constancy is binocular rivalry (Blake and Logothetis, 2002), where two different images are presented, one to each eye. Under suitable conditions, the percept of the subject flips between the two images, making the subject see one image for a few seconds, and then the other for a few seconds, and so on back and forth. NCC researchers have tried to look for neural activity that reflects this dynamic change of percept, in the confidence that such
activity could not be due to a change in the retinal input (Blake and Logothetis, 2002). If the retinal input is the same, what changes in the percept must takes place inside our mind, inside our consciousness - so the logic goes. By mapping these changes we hope to uncover the NCC.

Amid converging neuroimaging evidence, a consensus has now emerged that activity in a "fronto-parietal network" is an important component of the NCC (Rees et al., 2002), in addition to activity in modality specific areas (such as occipital and temporal areas for visual stimuli). It has been proposed that neurons in the prefrontal and parietal cortices form a global neuronal workspace for consciousness (Dehaene et al., 2003). When information from any modality enters this workspace, it becomes conscious (Baars, 1988). The emergence of this consensus could be considered a major achievement of neuroimaging research, which allows us to monitor neural activity across the whole brain. Prior to the availability of this method, studies of neuroanatomy and cortical lesions have led to the distinction between the dorsal (upper) and the ventral (lower) stream in visual information processing (Ungerleider and Haxby, 1994). Based on studies of patients with lesions to either the parietal or temporal cortex, it has been suggested that the dorsal stream, which mainly routes through the superior parietal cortex, is important for the online control of action, and that the processing is not necessarily conscious (Goodale and Milner, 1992). According to this influential view, visual consciousness depends more on the ventral stream, which mainly routes through the temporal cortex. Converging evidence from whole-brain neuroimaging has begun to undermine this view, highlighting the additional contributions of dorsal areas to visual consciousness.
However, it is important that we understand how specific these activations in the "fronto-parietal network" are. Tasks that are difficult, or simply require high attention, typically activate these regions (Hon et al., 2006; Duncan and Owen, 2000; Cabeza and Nyberg, 2000). Reaction times also correlate with activity in these regions (e.g. Gobel et al., 2004). It has also been reported that activity in this "fronto-parietal network" reflects the chewing of gum (Takada and Miyamoto, 2004); there was a significant difference in activity when subjects actual chewed gum, as compared to pseudo-chewing (without gum), which was supposed to control for jaw movements. Given that such a wide variety of tasks could activate this "network", which vaguely maps to a large part of the upper half of the brain, extra caution should be exercised in interpreting its supposed relationship with consciousness. In particular, although NCC researchers are extremely careful in controlling for the stimulus confound, it might still be possible that the results could be explained by other experimental confounds.

Given the role of the dorsal areas in action, the potential significance of a confound in performance should not be overlooked. In studying early visual areas which receive information from the eyes (via the thalamus), controlling for the visual stimulus confound is clearly important. However, in neuroimaging we often investigate neural activity from the whole brain, including higher cognitive areas. This means that merely controlling for stimulus constancy is not necessarily enough. Here I argue that performance, i.e. the ability to detect or discriminate visual targets, as measured by accuracy or reaction times in forced-choice tasks, is a critical confound in NCC research. Many findings in the NCC project could potentially reflect differences in performance, rather than perceptual consciousness itself.
For example, in using visual masking to render words invisible, we often give subjects a forced-choice discrimination task to make sure that the mask is working properly. To claim that the words in the masked condition are really invisible, we may try to demonstrate that subjects can only distinguish the words from meaningless letter strings at chance level, i.e. 50% correct. In the unmasked conditions, on the other hand, we argue that the words are visible because subjects can perform the discrimination at nearly 100% correct. When we compare the neural activity underlying the two conditions, how do we know that the results reflect a difference in consciousness of the words rather than the difference in the performance (50% vs 100% correct)? Similarly in a change blindness experiment, we compare trials where subjects see a change (hits) with the trials where subject do no see the change (misses). This means that we are essentially comparing 100% correct performance (hits) against 0% correct performance (misses).

One might think that binocular rivalry does not suffer from this problem, because subjects were not required to perform a task (apart from reporting their perceptual changes in some experiments; sometimes they are not even required to do so). However, it has been demonstrated that when subjects are asked to detect a small physical change in the presented images (such as an increase of contrast, or inception of a dot), they are more likely to be able to detect the change in the image they are seeing at the moment than the change in the image they do not see (Blake and Camisa, 1979). In other words, visual consciousness correlates with performance; conscious perception of an image predicts superior performance in change detection. Therefore, when we look for a change in activity that reflects the change in percept, it could be argued that it reflects the changes in detection performance instead.
One could argue that this is only potential performance, as we usually do not give subjects a detection task in binocular rivalry. But this is the same as arguing that we can eliminate a confound by not measuring it. Performance here does not refer to the actual motor action. It refers to the ability to detect or discriminate visual targets, or in other words, the effectiveness of visual information processing. These are what performance indexes such as reaction times or accuracies are meant to measure. Even if we do not measure the difference, it is there. And it could be the explanation of why we find a difference in neural activity.

3. Task Performance as a Measure of Consciousness

One may respond to the foregoing criticisms that the effectiveness of information processing is the same thing as consciousness. The argument would be that high performance in a psychophysical task is the hallmark of, or even one and the same as, perceptual consciousness. If one is conscious of the stimuli, one can perform well accordingly. If one is unconscious of the stimuli, one performs at chance. I shall argue that these claims are not true. In psychophysical experiments, performance is usually measured by the subjects' report of, or best guesses about, the stimulus (what is it?); this is different from their report about their conscious experience (what do you see?). There is empirical evidence that the two can be dissociated, and this means that task performance cannot be one and the same as perceptual consciousness.

First, task performance can exceed the level of consciousness. This is dramatically demonstrated in blindsight patients (Weiskrantz, 1986; Weiskrantz, 1999). After a lesion to the primary visual area, these patients report a lack of perceptual
consciousness in the affected region of the visual field. However, when forced to
guess the identity or presence of certain stimuli, they can perform well above chance
level, sometimes to an impressive range of 80%-90% correct. If we take performance
as a measure of consciousness, we would have to say that blindsight subjects do not
lack visual consciousness in their affected visual field. But they insist that they do.
Even in very well controlled laboratory experiments, using painstaking
psychophysical measures, blindsight subjects show clear signs of disturbed visual
consciousness (Azzopardi and Cowey, 1997; Weiskrantz, 1986; Weiskrantz, 1999; Cowey, 2004).

Second, task performance may underestimate the level of consciousness. In a classic
experiment on iconic memory, Sperling (1960) briefly presented an array of 3 rows
of 4 letters to subjects. When asked to report all 3 rows of letters, subjects could only
report a fraction. In another condition, immediately after the disappearance of the
letters, they were cued to report letters from a specific row. In this condition they
could perform nearly perfectly well, whichever row they were cued to report. This
means that, somehow, at some point after the disappearing of the letters, the subjects
had access to all letters, at least for a limited time. One could argue that the images of
all letters had entered perceptual consciousness, which seems to be
phenomenologically plausible (Block, in press). According to this view, when the
letters were briefly presented, subjects saw all of them. When asked to report all 12
letters, the subjects only fail because of a failure of memory. One could generalize
this to other tasks. One could always consciously see something, but fail to report it
later due to a failure of memory. In general, if the reporting procedure in the task is
not optimal, such as that it takes too long, as is probably the case in the Sperling
(1960) experiment, or that attention is distracted, one might fail to report the full content of what has entered consciousness.

Finally, whereas consciousness is unified, task performance is not. We typically believe that we cannot be conscious and unconscious of the same stimulus at the same time. This is because consciousness gives one coherent, unified perspective. However, Marcel (1993) has asked subjects to perform a visual task using different reporting modalities, such as manual, verbal, and eye-blink responding. The different modalities gave different performance levels. For example, sometimes the subjects get the answer right using blinks of the eye, whereas they do not with a manual response. This finding is very counterintuitive, but it has recently been replicated by my colleagues (Oliver Hulme and Barrie Roulston, personal communications). One could argue that the results challenge the concept of the unity of consciousness, but a more plausible interpretation is that whichever the reporting modality, perceptual consciousness is the same. It is only that in a particular reporting modality, unconscious guessing may be more effective than in others. Task performance, especially in forced-choice settings (which psychophysicists favour), can reflect an unknown and varying degree of influence by unconscious guessing, and is therefore sometimes an inconsistent measure of consciousness, as demonstrated in this example.

To sum up, when considered as a measure of consciousness, task performance sometimes captures too much, sometimes too little, and sometimes gives contradictory results. Given that task performance is such a poor measure of consciousness, why do we still use it in many neuroimaging experiments?
4. Objective vs Subjective Measures

Task performance, in particular performance in forced-choice tasks, is favoured by many researchers because it seems to be a relatively objective measure. After all, it is one main goal of science to try to characterize things in an objective fashion. If we ask subjects whether they consciously see something or not, their answer might be biased by their personal strategy or their interpretation of the phrase ‘consciously see’. Liberal subjects may frequently say yes, even when they only see a hint of a shadow. Conservative subjects may frequently say no, until they see the target clearly and vividly. It is therefore not a bad idea to ask subjects the relatively objective question of what is presented instead. Even if they are not sure, we force them to make the best guess. This way we can hopefully exhaust all the information the subjects have acquired from the stimulus presentation. To further eliminate bias, signal detection theory is typically applied to the analysis of the data (Macmillan and Creelman, 1991). This results in an objective measure of stimulus detection/discrimination sensitivity known as d', which is independent from the subject's choice of criterion or strategy (i.e. liberal or conservative). In other words, d' objectively characterizes a system's capacity or effectiveness to capture the information regarding the stimulus. Using d' as a measure of performance is the gold standard in psychophysical experiments, including studies on consciousness.

However, because of how objective and straightforward the analysis is, signal detection theory could actually also be applied to characterizing the sensitivity of a photodiode and other similarly simple electronic devices. In fact, signal detection theory was partly developed for such purposes (Green and Swets, 1966). The
realization that task performance essentially reflects any system's basic information processing effectiveness allows for an alternative and trivializing interpretation of current NCC findings. Typically, when the neural activity for a condition where subjects consciously perceive a stimulus is compared against an 'unconscious' condition, there is more activity in many regions throughout the cortex (including the 'fronto-parietal network'). However, few have reported results for the reverse comparison. One way to look at the findings is to shift our focus away from the term "consciousness", but to simply consider the fact that in the 'conscious' condition the brain is processing a lot of information, whereas in the 'unconscious' condition there is little to be processed. This is reflected by the difference in objectively measured task performance levels, which captures the basic effectiveness of information processing in the brain. The frontal and parietal lobes contain the important association cortical areas, which are highly connected with each other as well as with many areas in the rest of the brain. It is not surprising that when the brain is processing information in a highly effective and sophisticated fashion, as compared to no processing at all, the frontal and parietal areas are employed.

Therefore, according to this alternative explanation, in studying the NCC what we are actually looking at is the basic mechanism of how the brain processes information and produces useful responses. A suitable analogy would be the comparison between a normal working computer that is receiving input and producing output in an effective and productive fashion, and a malfunctioning computer which is neither responding to inputs nor doing any useful work. It would not be surprising to find out that there is more electrical activity in the major components of the computer when it is working normally.
Is this what we are interested in looking at? I suspect that a good number of NCC researchers are prepared to give a blunt "yes" to this question. Consciousness, according to them, is the basic mechanism by which the brain processes information. So studying the latter is the same as studying the former. But the reason that there is so much interest in the NCC project is that we have the powerful intuition that consciousness is more than simple information processing. Consciousness, that is the phenomenal and qualitative character of perception, seems to be beyond our normal understanding of mechanical information processing systems. This is why consciousness is considered to be such a mystery. Indeed, it has even been suggested that to fully understand consciousness we need to revise our framework of cognitive science and even the foundation of physics (Chalmers, 1996). The 'hard problem' of explaining the subjective character of consciousness is the reason why many scientists like myself have entered the field in the first place. We would not like to concede that in trying to study consciousness objectively, we end up just studying the basic mechanism of information processing in the brain - for this is exactly what we suspect consciousness outstrips. At the very least, consciousness has to be a specific form of information processing that is not universal to all neural processes (Lau, in press).

This is not to say that the basic mechanism of neural information processing is unimportant to understanding consciousness. In fact, outside of the context of the NCC project, the term "consciousness" is often used to refer to our level of wakefulness (Laureys, 2005), and this is closely related to the basic capacity with which the brain can process information regarding the external world. Even if we keep our focus on perceptual consciousness, it is reasonable to assume that the brain
has to be able to process or represent the information regarding a stimulus in order for us to be conscious of it. But similarly, in the long run we would also need a healthy respiratory system for us to consciously see, assuming that dead people do not enjoy perceptual consciousness. The question here is not whether the task performance is relevant to the study of consciousness. But, given the fact that task performance is a poor measure of perceptual consciousness, the question is whether we should consider subjective reports as a more proper measure. After all, it is the subjective character of consciousness that makes the problem so intriguing. This choice between objective and subjective measures has to be made because these measures do not always agree.

One classic example of a subjective measure of consciousness is the commentary key introduced by Weiskrantz (1999). In studying blindsight patients, after a forced-choice question concerning the stimulus, a subjective question is often asked as to whether the patients consciously see the stimulus event or are just guessing. Given that the patients can perform well, though reporting that they are just guessing, one would sensibly consider the subjective report to be more reliable. When blindsight patients claim that they are guessing, we consider them to be perceptually unconscious even though their forced-choice performance is high.

Many have already argued, on theoretical grounds, that subjective measures are the more valid and relevant measures of consciousness (Dienes and Perner, 1994; Chalmers, 1994). Recently, new subjective measures are also being developed for NCC studies (Sergent and Dehaene, 2004; Sergent et al., 2005). The 'Seen/Guesses' report in the commentary key paradigm could be reduced to a single button press, and is not substantially more difficult to analyze than other psychophysical data. There are
concerns about personal bias in subjective measures, but this could circumvented by comparing the measure within the same subject in different conditions (Lau and Passingham, 2006). The undeniable merit of subjective measures is their face validity: they are reports about the conscious experience itself, rather than reports about something else, that is the stimulus.

One might think that the choice between subjective and objective measures is not so important, because the two correlate most of the time anyway. However, as we have seen in the examples from the last section, they do not always correlate. And precisely because they correlate most but not all of the time, we need to apply extra caution in disentangling the two, so as to find out whether our NCC results reflect one measure or the other.

Just acknowledging the importance of subjective measures is not enough. It is important to note we do not solve the problem of performance confound by simply collecting subjective reports as an additional procedure, or worse as an afterthought. Nor is it enough to just replace objective measures with subjective measures. In searching for the NCC, we need to take extra steps to ensure that any difference of neural activity in our comparison between conditions reflects the difference in subjective reports, but not the difference in performance. One way to achieve this is to collect data for both measures, and set up conditions where performance levels are matched, and try to look for a difference in the subject report.

5. Performance Matching
The idea of matching performance in studying perceptual consciousness is not new.
Weiskrantz et al (1995) have suggested that the studies of blindsight patients offer such opportunity, because performance level for the 'blind' visual field could be as high as that afforded by the unimpaired visual field, if we choose the stimulus strengths for the two differently and carefully. Here I present an example of how this can also be done in healthy subjects in neuroimaging experiments. I have presented similar stimuli to the same visual location to the same subject in the same experiment, under conditions in which performance levels are matched, but the degree of subjectively reported consciousness differed. I show that this is a useful and practical strategy for uncovering the NCC.

In the experiment (Lau and Passingham, 2006), subjects were presented with either a square or a diamond in every trial, and they had to indicate which was presented to them (forced-choice performance). This visual target, however, was masked by a metacontrast stimulus (Fig 1). Here the masking stimulus surrounds but does not overlap with the masked stimulus. One interesting property of metacontrast masking is that the masking function is u-shaped: as one increases the temporal gap between the target and the mask (the stimulus onset asynchrony, SOA), the forced-choice performance decreases and then later increases again (Fig 2). This means that there will always be pairs of SOA points at which the performance level is matched (33ms and 100ms in the data shown). As in the 'commentary key' paradigm, on each trial, after the forced-choice, we collected subjective reports of whether the subjects saw the identity of the target or just guessed what it was. This subjective measure of perceptual consciousness differed significantly when the SOAs were 33ms and 100ms, even though the performance levels were matched.
We then used fMRI (functional magnetic resonance imaging) to scan subjects while they were performing this task. For each subject we focused on the two SOAs for which performance was matched but the proportion of guesses differed. The higher level of subjectively reported perceptual consciousness in the longer SOA condition was associated with a higher intensity of neural activity in the mid-dorsolateral prefrontal cortex (mid-DLPFC). This region was the only area where we found a difference in the whole brain. In other words we did not find a difference in either parietal cortex or the ventral visual stream.

Our result differs substantially from the results of previous NCC studies, since typically these report widespread activity throughout the cortex (Rees et al., 2002). In the context of the debate as to whether the dorsal visual stream contributes to visual consciousness, it is important to note that the mid-DLPFC has been formally characterized as a final converging point of the dorsal and ventral visual streams (Young, 1992). I have elsewhere proposed that perceptual consciousness depends on the higher-order statistical learning of the behaviour of early sensory signals, and have suggested that the mid-DLPFC is likely to be playing precisely this role (Lau, in press).

**6. Closing Remarks**

The point of this chapter is not to argue for any particular notion of NCC. There are many alternative notions, such as 40hz synchrony (Singer, 2001) and recurrent processing (Lamme, 2003), and these have not been discussed. However, though these notions differ, the same formal argument can be made, that is that one must eliminate a performance confound when studying these too.
The main point here is to raise certain concerns about the behavioural paradigms we use in the search for the NCC. The arguments do not depend on the new data I presented. Rather they depend on more 'classic' findings such as blindsight, and seminal experiments by Sperling (1960) and Marcel (1993). These results, some reported well before the popularity of the NCC project, gave early warnings about the validity of using forced-choice task performance as a measure of consciousness. In particular it was the study of blindsight that led to the development of the "commentary key" paradigm, and to the important idea that we could study consciousness while having performance levels matched (Weiskrantz et al., 1995).

The fact that the foregoing arguments depend essentially on old findings suggests that the answer to the awkward question of "Are we studying consciousness yet?" could be equally awkward: "We had been! We were almost there!" Of course, as a neuroimager himself the author is by no means implying that neuroimaging has set our progress backwards. Without doubt neuroimaging has shed great light on the nature of consciousness and its associated neural mechanisms. However, we should not take for granted that powerful machines or sophisticated data analysis algorithms will guarantee conceptually interesting results. The dramatic progress in technology in neuroscience has given us a great sense of optimism, but this should not mean that we ignore important conceptual issues in experimental research.
Acknowledgements

I thank Chris Frith, Tom Schofield, Dick Passingham, Uriah Kriegel, Dave Chalmers, David Rosenthal, and the editors for reading earlier drafts of this chapter. This final version has benefited greatly from their valuable comments.

Reference


Figures

Figure 1 – Visual Discrimination Task With Metacontrast Masking

After the presentation of the target and the mask, the participants were first asked to decide whether a diamond or a square was presented. Then, they had to indicate whether they actually saw the target, or that they simply guessed the answer. Shown in the brackets are the durations of each stimulus.
Figure 2 – Consciousness and Performance

These results were obtained using the procedure described in figure 1, except that it also included trials where the mask was presented before the target (paracontrast masking). Note that at the SOAs where the performance levels (% correct) were the same (e.g. 33 ms and 100 ms), the awareness levels (% seen) differed significantly.