Rüdiger Vaas
Institute of Philosophy, University of Stuttgart
Institute of Zoology, University of Hohenheim
e-mail: Ruediger.Vaas@t-online.de

Why Neural Correlates Of Consciousness Are Fine, But Not Enough

Summary

The existence of neural correlates of consciousness (NCC) is not enough for philosophical purposes. On the other hand, there's more to NCC than meets the sceptic's eye.

(I) NCC are useful for a better understanding of conscious experience, for instance: (1) NCC are helpful to explain phenomenological features of consciousness – e.g., dreaming. (2) NCC can account for phenomenological opaque facts – e.g., the temporal structure of consciousness. (3) NCC reveal properties and functions of consciousness which cannot be elucidated either by introspective phenomenology or by psychological experiments alone – e.g., vision.

(II) There are crucial problems and shortcomings of NCC: (1) Correlation implies neither causation nor identity. (2) There are limitations of empirical access due to the problem of other minds and the problem of self-deception, and (3) due to the restrictions provided by inter- and intraindividual variations. (4) NCC cannot be caught by neuroscience alone because of the externalistic content of representations. Therefore, NCC are not sufficient for a naturalistic theory of mind, (5) nor are they necessary because of the possibility of multiple realization.

(III) Nevertheless, NCC are relevant and important for the mind-body problem: (1) NCC reveal features that are necessary at least for behavioral manifestations of human consciousness. (2) But NCC are compatible with very different proposals for a solution of the mind-body problem. This seems to be both advantageous and detrimental. (3) NCC restrict nomological identity accounts. (4) The investigation of NCC can refute empirical arguments for interactionism as a case study of John Eccles' dualistic proposals will show. (5) The discoveries of NCC cannot establish a naturalistic theory of mind alone, for which, e.g., a principle of supervenience and a further condition – and therefore philosophical arguments – are required.
If our brains were simple, we would be too simple to understand them. Mario Puzo

The understanding of smart brains and minds requires smart techniques. Within the last decade a rapid progress of monitoring brain activities with imaging techniques like positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), with high-resolution electro- and magnetoencephalography (EEG, MEG), event-related potentials, intracranial multiple-electrode recording and transcranial magnetic stimulation etc. led to a growing body of evidence that behavioral activities and mental events – as far as verbal reports and (neuro)psychological test batteries can tell – are accompanied by certain neural states (see, e.g., Gazzaniga, 1995, for an introduction). This might count as a support for naturalistic theories of mind, which argue that consciousness supervenes on or is identical with or is realized or caused by certain brain states and processes. Let's accept the existence of such neural correlates of consciousness, NCC for short (please note the plural). We may safely do so at least for the sake of the argument. There is no need for them to be strictly localized. It is sufficient if they take the form of more or less robust spatial and/or temporal distributed relations within the nervous system, especially the brain. Some of these relations may probably be described only in abstract ways of representations, for example, neuronal network theory, the vector space approach or by order parameters in the framework of synergetics (cf., e.g., Churchland, 1995, Flohr, 1995, Haken, 1996, and Jirsa and Vaas, 1995). And even within neurophysiology, we must refer to many different levels of organisation and complexity – at least from the properties of single molecules on the one hand to the sensomotoric foundations and control of behavior on the other. Unfortunately, even this rather weak account is not enough for the more far-reaching and ambitious scientific and philosophical purposes like the understanding of intentionality, phenomenal awareness (qualia), self- and I-consciousness. On the other hand, there's more to NCC than meets the sceptic's eye.

Here, I'd like to give a brief sketch of three philosophical issues connected with the scientific investigation of NCC. I'm afraid that it is possible to go into details only occasionally, because my main point is trying to survey this diverse field. First, I shall demonstrate how NCC are helpful to explain properties of our conscious experience contrary to some criticism of neurophilosophy. Second, I shall review some problems and shortcomings of NCC. Third, I shall illustrate the relevance and importance but also some limitations of NCC to the still rather notorious mind-body problem.

I. Neuronal And Phenomenal Features Of Consciousness

Despite the ostensible fundamental difference between our subjective experience and the third person perspective of science, NCC can address or even explain properties of consciousness. Let me show three examples for this:

I.1. "Sweet Dreams Are Made Of This" – NCC Meet Phenomenology

NCC help us to explain phenomenal features of consciousness. An interesting example is the neuropsychology of REM (rapid eye movement) sleep dreaming (for the following see, e.g., Braun et al., 1998, Hobson, Stickgold, and Place-Schott, 1998, Maquet et al., 1996, Solms, 1997, and Vaas, 1998). In recent times PET studies have unveiled which brain areas show an increased or decreased metabolism during dreaming. Of philosophical relevance here is that these findings provide a deeper understanding of some well-known phenomenal features of dreaming which consist in a linkage between already known NCC within the awake brain on the one hand and the recently discovered neural correlates of phenomenal features during dreaming on the other.

This linkage is established in the following way:

- At first we know an already reliable fact of crude NCC, namely that a specific conscious state in the awake person goes along with a specific brain activity.
- Furthermore, we are familiar with some features of consciousness during REM sleep dreaming from our first-person perspective (and from the reports of others).
• Hence we can predict the neural correlates of these conscious states of dreaming, assuming that there is a connection with the NCC of similar awake states. If this prediction turns out to be true, we can reasonably assume that these underlying brain states somehow determine the features of REM sleep dreaming. At least we have stronger reasons to believe that there are lawful connections between specific conscious states and neural states.

Table 1: Neuropsychology of dreaming. ↑ = increase of metabolism, ↓ = decrease.

<table>
<thead>
<tr>
<th>Phenomenal Features of REM Sleep Dreaming</th>
<th>Neurophysiological Correlates</th>
</tr>
</thead>
<tbody>
<tr>
<td>vivid visual hallucinosis</td>
<td>extrastriate cortices ↑</td>
</tr>
<tr>
<td>spatial imagery construction</td>
<td>primary visual cortex ↓</td>
</tr>
<tr>
<td>motoric hallucinosis</td>
<td>right parietal operculum ↓</td>
</tr>
<tr>
<td>bizarreness (incongruity, discontinuity, uncertainty)</td>
<td>basal ganglia</td>
</tr>
<tr>
<td>delusion (being duped into believing to be awake)</td>
<td>frontal cortex (dorsal, orbital) ↓</td>
</tr>
<tr>
<td>deficits of self-reflective awareness, directed thought, insight in illogical and impossible experience, and memory (“dream amnesia”)</td>
<td>aminergic demodulation (noradrenergic and serotoninergic neurotransmitter ↓)</td>
</tr>
<tr>
<td>strong emotions (especially anxiety, fear, anger, elation)</td>
<td>(para)limbic system ↑</td>
</tr>
<tr>
<td></td>
<td>amygdala ↑</td>
</tr>
<tr>
<td></td>
<td>anterior cingulate ↑</td>
</tr>
<tr>
<td></td>
<td>temporal pole ↑</td>
</tr>
</tbody>
</table>

The latest PET studies provide a consistent and coherent picture indeed. Table 1 gives some examples. It lists the well-known phenomenal features of dreaming and the recently discovered neurophysiological correlates. Their anatomical locations were already known as necessary for the summarized mental features or their undisturbed functioning of awake persons. These studies led to the hypothesis that dreaming is a result of (1) a functionally isolated loop between the extrastriate cortices and the (para)limbic system including the amygdala, largely disconnected from sensory input and motor output due to inhibition of the striatum and frontal cortex respectively, (2) a chaotic autoactivation process within the extrastriate cortices and the (para)limbic system, triggered by an increased input from the brainstem and the basal forebrain, and (3) the absence of top-down control because of an inhibited frontal cortex. The effects of brain lesions on dreaming strengthen this proposal. Lesions which are correlated with deficits or accentuations of dream experience remarkably complement the studies of brain metabolism. For example, PET studies revealed an increase of activation of pontine tegmentum, limbic structures, extrastriate cortex and right parietal operculum in REM sleep while lesions of these areas decrease dreaming or have no effects; and a decrease of metabolism in striate cortex and dorsolateral prefrontal cortex measured by PET correspond to no changes of dreaming after lesions of these areas.

Of course, PET studies provide only an indirect, approximate insight into NCC with spatial and temporal resolutions of a few millimetres and seconds, averaged over many experiments or individuals. But they give us important information about the functional architecture of our brains and demonstrate that different aspects of our mental life depend on different areas of the brain. And there are other, more sensitive methods like EEG, MEG and psychophysical experiments that unveil temporal constraints of conscious experience, which sometimes contradict what we find by introspection, or are at least not accessible in this way. This leads me to my next point.

I.2. “The Times, They Are A Changing” – NCC On Stage

NCC can account for phenomenologically opaque facts – the temporal structure of consciousness, for instance (see Dennett and Kinsbourne, 1992, Metzinger, 1995, and Vaas, 1997). The phenomenal temporal continuity and homogeneity is in a sense an illusion, because there are discrete "temporal units of consciousness", e.g., a succession of temporal windows of conscious perception and action, each about two to three seconds long, and a succession of
perceptual-motor thresholds in the order of 30 to 40 milliseconds (Pöppel, 1985). There is even a sophisticated confusion of the objective temporal order.

Take the phenomenon of the "cutaneous rabbit", as Daniel Dennett (1991, p. 142 f.) called it: The subject's arm rests cushioned on a table, and mechanical tappers are placed at two or three locations along the arm, up to a foot apart. A rhythmic series of taps is delivered by the tappers, e.g. five at the wrist followed by two near the elbow and three more on the upper arm. The taps are delivered in intervals between 50 and 200 milliseconds. The astonishing effect is that the subject feels the taps traveling in regular sequence over equidistant points up the arm – as if a little animal were hopping along it. But the brain obviously cannot "know" about a tap at the elbow until after it happens. So there is a change in the representation of the intervals, a mistaken interpretation after the taps are registered.

Another example are two alternately flashing lights in a dark room, separated by a distance not too large, which are perceived as one single, rapid moving dot. If the lights have different colors, it seems to the subject that there is a change in color in the middle of the move (although the second light is not yet shining at this very moment!). If one is not willing to believe in clairvoyance, we must assume that our brain can predate this flashing somewhat into the past.

An even more surprising discovery were Benjamin Libet's (1993) experimental findings of neural delays, retrograde stimulus masking and subjective referral backwards in time. The comparison between an electrical stimulus applied to the hand and a stimulus applied directly to the corresponding cortex area demonstrates that we perceive the hand but not the cortex stimulus 0.5 seconds later. But we do not realize this delay, because our brain shifts it back into the past. Thus, there seems to be a double illusion: our experience of time is, contrary to common sense, pretty much behind the events and we are not able to realize this because our temporal frame of reference is also shifted. Furthermore, some informations could be masked by later informations. In addition, Libet and others have also demonstrated that unconscious brain processes (marked by a "readiness potential") occur at least 0.35 seconds before the conscious intention to act, although it seems to us that our intention causes our action.

These results are subject to controversial interpretations, to be sure (e.g., Churchland, 1981, and Gomes, 1998). But they suggest that what is objectively determinable as an unequivocal order of before-and-after need not respond necessarily and unequivocally to our subjective experience. The spatiotemporal representation of some stimuli seems to be a construct of our neural activities and need not correspond to the real order (just as it is pretty well documented that we remove our hand from the fire before we feel the pain, contrary to our subjective experience). There is no need for more precision as long as these constructs are not biologically disadvantageous, sorted out by the sieves of evolution. Furthermore, there is no neuroanatomical center where everything comes together. There is no "Cartesian Theatre", as Dennett (1991) has called this view, which underlies our somewhat misguided conception of temporal order and leads to an infinite regress (the well-known homunculus problem). There is a symmetry of distributed, parallel-processed representations or "multiple drafts" instead, interacting with each other, organizing themselves in a complex, nonlinear way and bound together by spatial and temporal patterns, guided perhaps by far-reaching synchronized neural activities (Metzinger, 1995, and Singer, 1996).

I.3. "Paint It Black" Or "In The White Room" – NCC And A Vision Of The Brain

NCC reveal properties and functions of consciousness which cannot be elucidated either by introspective phenomenology or by psychological experiments alone – vision, for instance (cf. Zeki, 1993, 1997). Here, the studies of lesions due to stroke, ischaemia, injury, tumor, carbon monoxide poisoning and so on are very instructive. They give important insights in the neural organisation underlying consciousness.

• For instance, for a person with cerebral achromatopsia due to a lesion of V4 in the fusiform gyrus it is sometimes not only impossible to perceive colors anymore but also to imagine, remember and dream them. (In some cases color imagery is preserved in achromatopsics probably due to an afferent disconnection of visual input to stored color representations i.e. an impairment of bottom-up but not top-down processing.)

• Bilateral damage of the parieto-occipital lobe (the "where-system") leads to an inability to localize visual stimuli in space and to accurately describe the location of familiar objects or landmarks from memory,
whereas bilateral damage of the inferior temporal lobe (the "what-system") causes impairment of perceiving object identity from appearance and describing object appearance from memory.

Furthermore, some patients with lesions in V1 but spared subcortical pathways (lateral geniculate nucleus and pulvinar nucleus; the latter gets inputs from the retina via the superior colliculus in the midbrain), are able to perceive motion in their otherwise blind fields (but they do not see stationary objects). In contrast to "blindsight", where subjects are able to discriminate visual stimuli which they are not consciously aware of having seen, patients are conscious of residual motion vision without being forced to guess whether there is motion or not. They describe this experience, e.g., as "vague and shadowy", similar to the experience of a normal subject who shuts his eyes, looks out of the window, moves his hand in front of him and sees the shadows. Some patients can only detect fast-moving stimuli, whereas others can only detect very slowly moving ones.

These examples suggest, as Semir Zeki (1997, p. 143) has pointed out, that "there may be many more or less separate consciousnesses for different attributes at least of the visual world, based on activities in separate visual areas." Thus, at least visual consciousness is in some sense modular and not exclusively dependent either upon a single cortical area (or multiple, but intimately connected areas) or upon the healthy functioning of the entire system. Despite our phenomenal impression that our consciousness is indivisible and maybe even not spatial, lesions show that it is possible to lose consciousness not only completely, but also in bits and pieces. Thus, consciousness has aspects of divisibility and spatiality even if we cannot recognize them by introspection.

II. Problems And Shortcomings

There are some crucial problems and limitations for NCC research, partially related with philosophical issues. This is one reason why metatheoretical reflections are required, which are part of what could be called neurophilosophy (Churchland, 1986).

II.1. "Smoke On The Water" Or "Riders On The Storm" – Limitations Of Correlations

Correlation implies neither causation nor identity (although it might be interpreted as an indication for one of them). For example, a strict correlation of birth rates and the size of stork populations does not mean that babies are made or brought by storks; and a strict correlation between the movements of soccer-players on a TV screen and in the stadium from which there is a live broadcast does not mean that the pictures on TV are identical with the soccer players (Zoglauer 1998, p. 106). Thus, NCC alone are not sufficient to prove that our conscious experience is caused by or identical with neural events. There is even a tension between the ways of talking about causation and identity here. But of course this depends strongly on our notions of causality and identity, which I cannot discuss in this context. In any case, neural events need not necessarily be identified with mental events or cause them. The correlations are compatible with very different other dependencies and ontologies (see III.5.), e.g., there could be a common cause in the past for both neural and mental states which are otherwise independent from each other (e.g., Gottfried Wilhelm Leibniz' preestablished harmony), or a continued intervention of a causal agent, e.g. God, synchronizing neural and mental states (e.g., occasionalism, as it was proposed by Arnold Geulinx and Nicolas Malebranche), or both states are two aspects of one and the same underlying microprocess, or the neural states are caused by the mental states and not vice versa, or the correlations are just an improbable or unexplainable coincidence. Furthermore, correlation does not imply identity because the simultaneous occurrence of a mental and a neural state is also consistent with parallelism and epiphenomenalism. It is not possible to refute a sufficiently sophisticated version of parallelism empirically, for instance by reference to NCC, because there is no way to distinguish between identity and parallelism empirically, as it was already admitted by Herbert Feigl (1958, p. 437 and 463).

II.2. "Private Investigations" Or "Empty Rooms" – Restricted Access For NCC

There are limitations of empirical access due to the well-known problem of subjectivity, the problem of other minds, and the problem of self-deception.
The problem of subjectivity, also called the "explanatory gap" or the "hard problem" (cf. Bieri, 1992, Chalmers, 1996, Jackson, 1982, Levine, 1983, and Nagel, 1974, 1980), is based on the dichotomy between the – in some sense – irreducible subjective character of our phenomenal experience (qualia) and the quasi-objectivistic explanations of science or any third-person description. We will never know "what it is like to be a bat" (or even my own twin brother) even if we could know everything physical including all NCC in every detail. But this is not necessarily an ontological problem and therefore a dread for naturalism (taken as an ontology). The difference between the first- and the third-person perspectivity might be just a result of the different modes of presentation and the different concepts which we apply, hence it could be seen as an epistemic and conceptual difference, but not an ontological one (cf. Tye, 1995).

The problem of other minds is obviously insurmountable: I cannot inspect the mind of other individuals, because I cannot log in their brains or minds or be part of them. I can only interpret their behavioral and maybe physiological actions and responses (including verbal reports). If they are zombies without inner experiences I would never know. If they are perfect actors or liars with quite different experiences I could not know either. Maybe I could detect contradictory physiological states, but to interpret them as contradictory, I still have to compare them with some standard and need an argument why this very case is not just an exception to the rule.

Furthermore, maybe it is even self-deceiving to attribute something like "inner experiences" to others. Maybe this is just an advantageous "intentional stance" to cope with their complex behavior or some sort of a linguistic illusion or a social construct (cf. Blackmore, 1999). Maybe it is self-deceptive to attribute "inner experiences" to myself, too. Studies of human development (which are admittedly controversial) give at least some hints that knowledge of one's own mental states is – contrary to common sense – not immediately given but as indirect as the knowledge of other minds: the infant has to learn at first by inference from its own behavior and the behavior of others in which mental state it is (Gopnik, 1993). Furthermore, there is reason to believe that we are systems which permanently confuse themselves with their own self-model, as Thomas Metzinger (1993) has put it. In doing this, we generate an ego-illusion, which is stable, coherent, and cannot be transcended subjectively, i.e. on the level of conscious experience itself.

But I do not want to propose that consciousness is something which is nothing at all. For in this case we need not look out for NCC. Nevertheless, as the discussion about the temporal structure of conscious events has shown, there is reason to believe that our mental states are not always transparent to ourself. How could we know if the correlations between neural and conscious events we observe are not biased in a systematic, misleading way? This is, for example, one of the main problems of Libet's interpretation of the temporal order: it is very difficult to prove that the person's clock-reading and self-reports are really reliable.

II.3. "Leave Mine To Me" – Variations And Complexity

Another limitation of the detection of NCC are interindividual differences in neural structure and dynamics and intraindividual variations in the course of time. Not even identical twins have identical brains. Individual variations are one of the most important restrictions of a generalization of functional brain-imaging techniques. Furthermore, they undermine a fine-grained lawful one-to-one correspondence between neural and mental states: If it is possible at all that two persons are exactly in the same mental state at a given moment, this mental state obviously cannot depend on an exactly identical neural state even if we could neglect the environment (which of course we cannot, cf. II.4.). Thus, conscious states are realizable in different ways (cf. II.5.).

Interindividual differences could be enormous in cases of developmental impairments and plastic reactions after that. For instance, there are quite normal people who are missing big parts of their brain due to early childhood surgery. And in the case of hydrocephalus, an increase in the volume of the cerebrospinal fluid could stretch the cortex into a sheet of tissue only a few millimetres thick without damaging it and without impairing intelligence. In a computer tomography scan, the head of a hydrocephalus appears largely brainless.

Investigating the brain on the cellular and subcellular level, it is nearly impossible to get detailed maps of the neural networks and their activities and to handle their complexity. This is a major obstacle for the discovery of sufficiently precise NCC. Furthermore, it is an open question, which kind of reduction is possible (Vaas, 1995a), and which levels of description, abstraction and generalization are most useful for philosophical purposes.
II.4. "You Make Me Real" – Externalism For A Wider Perspective

Add to this that NCC cannot be caught by neuroscience alone because of the externalist content of representations (cf., e.g., Burge, 1986, Davidson, 1993, Davies, 1991, and Dretske, 1995) – or at least because of some externalistic components. Therefore, even identical brain states could correspond to different mental states if the environment was different. One might say consciousness is not in the head, but this seems to be throwing the baby out with the bathwater. It is enough to recognize that consciousness is not only in the head (or depends not only on it). And that is because the brain (and mind) is not a closed system. Otherwise reference would be impossible and I would be trapped in total isolation or even had to assume that solipsism is true. Mental states heavily depend on information from the body, its ontogeny and phylogeny, its present and past physical and social environment. There is permanent interaction of an organism with its surroundings (Hurley, 1998), and it is not possible to individuate conscious states without taking these (past or present) interactions at least implicitly or tacitly into consideration. Even if I could study my own NCC with an autocerebroscope (Feigl, 1958, p. 456) together with introspection, by-passing the problem of other minds (II.2.), my mental contents would still be dependent on external references including a public language. Hence, given that a form of externalism is true, consciousness cannot be reduced to brain states alone. (But of course this does not mean that it could not be ontologically reduced to physical states in general – externalism is a form of naturalism with a wider scope, cf. III.5.) Therefore, NCC are not sufficient for a naturalistic theory of mind.

II.5. "The Soft Parade" – Nonneuronal Intelligence

Nor are they necessary, because of the possible existence of conscious artificial and extraterrestrial intelligences. Thus, due to its presumed multiple realizability (cf. Putnam, 1975), consciousness should be viewed not as a product of neural tissues only. This is of course no limitation of NCC research, though it restricts philosophical generalizations about the nature of consciousness. However, although multiple realizability may well rule out a general, uniform mind-matter reduction, it entails the possibility of locally reducing mental states to physical ones – and perhaps this is all the reduction we need or could want (cf. Kim, 1996, p. 234). Explanations, and hence reductions, are domain relative.

III. Neural Correlates And The Mind-Matter Problem

Despite the problems and shortcomings of NCC I have just reviewed, NCC are of some relevance and importance for the still rather notorious mind-body or mind-brain problem (cf., e.g., Kim, 1996, and Rey, 1997, for an introduction), which could be more precisely called the mind-matter problem, and for various philosophical issues, e.g., the nature of perception, representation, decision-making, action, consciousness and self-awareness. Neuroscience, among other scientific disciplines, is now able to suggest experimentally constrained hypotheses of philosophical relevance. Philosophers cannot speculate fruitfully on these issues in ignorance of the data. On the other hand, neuroscience becomes necessarily more and more involved with philosophical issues. Therefore, a transdisciplinary teamwork is required (Churchland, 1986, 1996).

III.1. "Piece Of Mind" – Necessarily Human

In spite of not being necessary for consciousness in general – because of the multiple realizability of consciousness – NCC reveal features that are necessary at least for behavioral manifestations of human (or primate or vertebrate?) consciousness and self-awareness (cf. Vaas, 1996). This can easily be learned from lesion studies, coma, stupor, anaesthesia and the reversible breakdown of NCC due to epilepsy, transcranial magnetic stimulation or simply dreamless sleep. Furthermore, as the neuropsychology of vision impairments shows (I.3.), there are even specialized, i.e., somewhat modular brain areas whose destruction leads to the selective loss of conscious experiences, color awareness, for example.

NCC are compatible with very different proposals for a solution of the mind-brain problem. The existence of NCC can and does not prove that a version of the mind-brain identity theory is right, although this existence is perfectly in harmony with it. But classical dualists would still argue that the correlates are just a consequence or a sign of the brain's interaction with an ontologically independent mind as renowned neuroscientists like Charles Sherrington, Wilder Penfield and John Eccles thought. A parallelism without interactions of neuronal and conscious events in the style of Leibniz is still possible, too, and it is even favoured in a somewhat more down-to-earth fashion by neuropsychologists like Detlev Linke and Martin Kurthen (1988). And epiphenomenalism, which holds only a one-way influence from matter to mind but not vice versa, is not falsified by the discoveries of NCC, either. Thus, one may subscribe to the existence of NCC and can still postulate that one dualism or another is true. This seems to be both advantageous (for dualists) and detrimental (for naturalists).

Of course these issues depend on our notion of causality. If for example interactionism is true, causality cannot solely be restricted to physical events like transfer of energy or momentum. On the contrary, a dualist has to postulate the existence of (up to) four qualitatively different cause-and-effect relations: interactions between physical events, interactions between mental events and, furthermore, physical-mental and mental-physical interactions. Parallelism requires the existence of two distinct sorts of causal relations (physical-physical and mental-mental interactions), so does epiphenomenalism (physical-physical and physical-mental interactions). Naturalism and, maybe, idealism, needs just one. If we subscribe to a Humean account of causality instead, a description of correlations is all there is. Therefore, the existence of NCC would be compatible with most of the positions regarding the mind-brain problem.

Please note: I do not say that all these positions are equally plausible. My claim is only that the existence of NCC alone cannot prove or disprove any of these positions. Thus, philosophical arguments are needed, and this is just one reason for the importance of neurophilosophy.

III.3. "Shades Of Truth" – Brain Imaging Is Not Mind-Reading

The existence of NCC restricts nomological and identity accounts for the mind-brain problem because of the already mentioned inter- and intraindividual variations and the externalist content of representations. We cannot strictly define identity conditions except by means of potentially infinite – or at least impracticably many – conjunctions. That is why the modern brain imaging techniques only show some more or less crude categories or types of consciousness (or contents of consciousness) like logical reasoning in dorsal frontal cortex, different sensory modes for example in different parts of the occipital, temporal and parietal cortices, or memories of different syntactic forms in the temporal lobe – but not the actual content of the logical reasoning or the sensory experience or the words. Thus, strictly speaking, brain-reading will never turn into mind-reading.

Even if we could scan all the brain structures and neural events of an individual with any desired resolution and without destroying it (which seems to be practically impossible), we would still not know what kind of experience or thought this individual exactly had "in mind" at a given moment of time. And this is not only true because we were missing his or her environment and the relevant past and we cannot falsify strange possibilities like inverted or absent qualia, but also because of the chaotic, i.e., nonlinear dynamics of the neural events and the potentially holistic nature of mental contents.

But these restrictions do not imply NCC research is useless or irrelevant for philosophical reasoning. On the contrary, the investigation of NCC can strengthen or refute empirical arguments for one or the other position in our discussion about the relation between brain and mind and the nature of consciousness. For example, let's take the interactionist dualism of the late Sir John Eccles, who was a winner of the Nobel prize for physiology in 1963, as a case study.

III.4. "Oh, Hello Mr. Soul, I Dropped By To Pick Up A Reason" – A Case Study

Eccles has started brain research because of religious reasons (Eccles, 1994a, p. 13, Popper and Eccles, 1977, p. 357) and presented probably the most elaborated neurophilosophical proposal for a place of interactionism in a Cartesian tradition, believing that mind is "independent" and "autonomous" from the world of matter-energy (Eccles, 1994a, p. 102 and 80). He has developed his speculations on "the Self and its brain" from 1951 until recent years and wanted "to challenge
and negate materialism and to reinstate the spiritual self as the controller of the brain” (Eccles, 1994a, p. X). It's important to recognize that he understood his account "as a hypothesis in Popperian scientific method" (Eccles, 1994a, p. IX; cf. Popper and Eccles, 1977, p. 375), and even saw "empirical evidence" for dualism (Eccles, 1987, p. 295 ff.); therefore Eccles has to allow (and reckon with) the possibility of an empirical refutation (Vaas, 1997b).

For Eccles, only minute fractions of the material world are associated with mental states (subjective experiences): the liaison-brains. That is where the conscious mind grasps the neuronal activities like a searchlight, selects and modulates what it is interested in and integrates it to a unified experience. Eccles believes that there are special modules which are linked to the mind like radio transmitters and receivers (Popper and Eccles, 1997, ch. E7). These modules consist of what he called dendrons. A dendron is a composite made of the bunching of the apical dendrites of pyramidal cells found especially in neocortical lamina V. According to Eccles (1994a, p. 136), "each dendron is linked with a psychon, giving its own characteristic unitary experience". There are "forty million psychons for an estimated forty million dendrons of the human brain" (p. 88 and 98). So all mental events and experiences "are a composite of elemental or unitary mental events"; each of these psychons is reciprocally linked in some unique manner to its dendron" (p. 87).

In 1992, Eccles (1994a, ch. 9) tried to explain the mechanism of mind-brain interaction – for him: the interplay between psychons and dendrons – on the basis of quantum effects. The mind becomes neurally effective by momentarily increasing the probabilities for exocytosis in a whole dendron without violating physical laws of conservation. The essential place of this interaction "is at individual microsites, the presynaptic vesicular grids of the boutons" (p. 82). Mental influences "do no more than alter the probability of emission of a vesicle already in apposition" (p. 73); they change synaptic activities on the quantum mechanical level without violating the conservation of energy or momentum and are able to increase (the probability of) exocytosis by quantum tunneling. This holds not only for intentional acts, voluntary commands for instance, but also for attentional acts: Attention can "activate any selected parts of the neocortex at will" and can increase the frequency of the impulse discharges in the pyramidal cells of a dendron (p. 174); this dendron triggers the excitation of its associated psychon to give an increased experience of a sensation, e.g., pain; "conversely, if attention is concentrated elsewhere, there will be less activation of the dendrons of the nociceptive cortex and pain will be alleviated" (Eccles, 1994b, p. 17). "Consciousness is experienced in the brain where you evoke it by your attention, which plays on selected areas of the cerebral cortex to give excitation. That excitation leads to amplified dendron responses to sensory inputs and so to psychon activations and consciousness. Superimposed on this simple attentional operation there would be a continuing dialogue between attention by the self and the selected neocortical areas with their sensory inputs" (Eccles, 1994a, p. 176).

Apart from philosophical problems of ontological interactionism in general and Eccles' dualism in particular, this proposal runs into empirical difficulties, too. Here are ten arguments based on empirical research against Eccles' empirical claims and ontological inferences.

(1) In earlier works, Eccles localized the "liaison-brain", where the Self should get into contact with matter, mainly in the dominant cortex hemisphere (endowed with language) (Popper and Eccles, 1977, p. 326).

• However, there are examples (like that of nine year old Alex, born with Sturge-Weber syndrome) of first language acquisition and remaining conscious after a complete ablation of the left hemisphere (Vargha-Khadem et al., 1997).

• Furthermore, PET studies of right-handed healthy people demonstrate clearly that autobiographic memories (thus, very intimate parts of Ecclesian Selves) are located in the nondominant hemisphere (increase of activities in the right temporal cortex, (para)hippocampus, posterior cingulate, insula, prefrontal cortex, decrease of various parts of the left cortex) (Fink et al., 1996).

• Eccles (e.g., 1987, 1994c) has speculated about the supplementary motor area (SMA) as the physical stage for interaction with the mind's voluntary commands – i.e., as the working place for free will. However, after temporary interruption of voluntary movements, there is an extensive recovery even after a complete ablation of SMA (Kolb and Whishaw, 1996, p. 262).

If on the other hand attention can "activate any selected part of the neocortex at will", as Eccles wrote more recently (1994a, p. 174, cf. p. 79), it is unclear why there are actually neuronal constraints of consciousness and volition and why the Self cannot sail around specific impairments of them.

(2) If the mind is in some ontological sense autonomous and independent of matter as interactionism assumes, it is difficult to see why specific brain lesions affect imagination as well as...
perception. However, mental imagery seems to be the efferent activation of some subset of the brain's visual areas, subserving the same types of functions (what, where, color, spatial attention, and so on) in imagery and in perception. As already mentioned (I.3.), impairments of specific regions of the cortex specialized for vision do not only affect the perception, but also the imagination, memory, and dreaming of colors, for example. If the mind is somewhat autonomous, how could that be? A similar reply can be made with respect to memory. Declarative memory is located in the neocortex, as it seems. If the mind can "read" informations from the neocortex and "write" them into it in some sense, it is mysterious why lesions in deeper structures (hippocampus, mammillary bodies, thalamic nuclei etc.) cause amnesia (cf. Vaas, 1994).

(3) Exocytosis depends on large proteins e.g. the formation of fusion pores (cf. Walch-Solimena, Jahn and Südhof, 1993, Weis and Scheller, 1998). It is unlikely that quantum effects can play a significant role here i.e. trigger configuration changes on the level of macromolecules. Under the known presynaptic conditions, Heisenberg's Uncertainty Principle $\Delta p \times \Delta x \geq \hbar$ works only for masses in the range of hydrogen atoms and for time scales in the femtosecond regime. It is by no means clear how such tiny quantum effects could trigger exocytosis i.e. could open the presynaptic and vesicle membrane. Even worse, if they really could, it is a mystery why these acausal events do not disrupt the neuronal activities and hence the organization of perception and behavior.

(4) There is a tension between mental causation and quantum physical indeterminism, i.e., acausality. Of course, Eccles needs the indeterministic and probabilistic nature of quantum effects to avoid violation of conservation laws. (For the sake of argument, the indeterministic nature of the quantum world might be taken for granted here despite some other views still under discussion.) This loophole in the causal nexus of nature is necessary but not sufficient for interactionist dualism. According to Eccles, the Self is able to change quantum probabilities (or select specific quantum states) and trigger neuronal activities in a goal-directed way. So not only must he fill causal gaps in nature with mindful interventions but he also needs some sort of nonphysical causal power — i.e., a new ontological type of causality — to explain the occurrence and efficacy of (self-) consciousness. But this implies a violation of the quantum mechanical probability distributions which is purely statistical.

(5) Furthermore, there is the problem of physical laws of conservation despite Eccles' claims that the mind does not violate the first law of thermodynamics. In the interactionistic picture, the mind must exchange information with the brain, but the current state of physics either postulates a transformation of matter or energy along with information processing or these events do not carry information at all. But even if such intended quantum effects do occur and do not violate conservation of energy because of Heisenberg's uncertainty principle $\Delta E \times \Delta t \geq \hbar$ (and would not decrease global entropy!), cortex activities depend as a matter of fact on large amounts of energy consumption, even during imagery and "pure" ideation, as PET and fMRI scans show. So why should and how could tiny quantum effects trigger the highest brain functions without any energy at all?

(6) Even if they could, how does an Ecclesian Self manage to control the myriads of transmitter-releasing synaptic vesicles without totally disrupting or disorganizing perception, thinking and motor commands? This is the well-known problem of Jordan's amplification hypothesis (Bünning 1935, 1943). (Inspired by quantum mechanics, Pascal Jordan (1934, 1938) has developed a similar hypothesis to rescue free will earlier this century.) And Eccles must postulate amplification (and explicitly does) because otherwise his hypothetical quantum effects would be ineffective.

(7) Next, in calculating the amplification effects very precisely to avoid catastrophic disorganization and errors in perception and behavior, the Self must know more than even quantum mechanics allows (because of Heisenberg's uncertainty principle). And it must be a little Laplacean demon, computing infinitely fast, because the brain is a complex system showing strong nonlinear (or chaotic) dynamics at many different levels which cannot be predicted in practice (Jirsa and Vaas, 1995, Vaas, 1995a).

(8) Furthermore, to give a dualistic explanation of Libet's (1993) experimental findings of neural delays, retrograde stimulus masking, and subjective referral backwards in time, Eccles needs to postulate that the Self "plays tricks with time" (Popper and Eccles, 1977, p. 364, 475), i.e., changes the temporal order! This violates causality and the direction of time. Rejecting his dualism and the (materialist) assumption of a "Cartesian theatre" where all perceptions gather and become conscious, one can explain these findings much better in terms of a "multiple draft" approach (Dennett, 1991), assuming that perceptual discriminations are distributed across the brain in both space and time (cf. I.2.).
(9) Electrophysiological results, demonstrating that unconscious brain processes (marked by a "readiness potential") occur at least 0.35 seconds before the conscious intention to act, were first interpreted by Eccles (1985, and in Eccles and Creutzfeldt, 1990, p. 207) as preconditions for the mind which grasps them, as an "incubation time of the selfconscious mind" (up to 0.8 seconds long), and as neural correlates or consequences of voluntary commands not exclusively determined in a neural way (Eccles, 1978, p. 1149, Popper and Eccles, 1977, p. 365, 293, 285) – contrary to the subjective reports. Later, Eccles (1994a, 138, and 1994c, 220 f.) simply denied the readiness potential without arguments or experimental data as an "artefact produced by the averaging technique of recording". However, the easiest explanation is that the readiness potential is the effect of neural input into the supplementary motor area from axons coming from other cortex areas (especially the prefrontal cortex) or the thalamus. Thus, there is no need for an intervention from another ontological realm (cf. Churchland, 1981).

(10) Finally, Eccles (1994a, p. 110) stipulates that the "transmission of psychon to psychon could explain the unity of our perceptions and of the inner world of our mind that we continually experience from moment to moment"; "it is the very nature of psychons to link together in providing a unified experience" (p. 136), the neural machinery cannot do it by itself (p. 22). But this is just a claim without argument and does not explain at all how sensory information was bound together by psychons. Even worse, it seems impossible to study such psychonic integrations in principle, because there is no way to break their connections if they aren't physical. So Eccles' suggestion, at least in this part, is not a scientific hypothesis at all. Neural accounts of the binding problem however are within experimental reach. Synchronous activities between neurons from different parts of the brain are promising candidates for a solution of the binding problem at least with regard to sensory integration (cf. Singer, 1996).

In conclusion, Eccles' neurophilosophical arguments fail to support his metaphysical speculations about the ontological difference between brain and mind and their neo-Cartesian interaction. Eccles cannot show by scientific reasoning what he claims.

III.5. "Wild Thing" Or "Stranger Than Fiction" – Toward A Naturalistic Theory Of Mind

The discovery of NCC cannot establish a naturalistic theory of mind alone because of the already mentioned conceptual, empirical and ontological reasons (see II., cf. also Vaas, 1995a). There is a variety of possible ontologies which are all consistent with the existence of NCC. This is not to say, however, that the discovery and study of NCC is irrelevant for consciousness research and philosophy of mind. It is always easy to construct weird, unfalsifiable metaphysical claims linking consciousness to some spooky angel dust, free-floating ghosts, sentient quantum particles or even blood pressure. As long as there is no empirical evidence, nobody needs to take such strange speculations seriously. It should be evident that the proponent of such speculations, not the opponent, has the burden of proof, because they are unfalsifiable. And it is quite wise to apply Ockham's razor here – i.e., the principle of ontological parsimony ("Pluralitas non est ponenda sine necessitate; essentia non sunt multiplicanda praeter necessitatem").

In any case, there are some more or less robust neurophysiological correlations of mental activities which provide a better understanding of these mental events (cf. I.). And a lession of lesions is that there are necessary neuropysiological conditions for the whole range of (human) expressions of mental events like seeing colors, remembering concepts or planning actions. These are necessary conditions at least if one is not willing to subscribe to idealism or to the view that the mind is still intact behind the stage of a damaged brain like an inaudible chatter behind a broken loud-speaker.

Because pointing to NCC is not sufficient for a naturalistic theory of mind, philosophical arguments are needed. Naturalism (sometimes – and somewhat misleadingly – also called physicalism or materialism) requires at least a form of supervenience and a further condition. Supervenience or a version of mind-brain identity theory alone is not sufficient.

An identity theory is not sufficient because it doesn't say what kind of ontological entity there is. If mental states are (identical with) physical states, one could still subscribe to the variety of idealism, ontologically reducing matter to mind via the identification and not vice versa. Even some neurophilosophers from the camp of radical constructivism seem to get (perhaps not intentionally) into the neighbourhood of solipsism which is a kind of idealism. An identity theory is also compatible with neutral monism, ontologically reducing mind and matter to a third "substance", for example, logos (Heraclitus), God (Baruch de Spinoza), elements (Ernst Mach), the absolute (Friedrich
Wilhelm Schelling), energy (Wilhelm Ostwald), sensibilia (Bertrand Russell) or an entangled indivisible unity of mental and physical aspects, for instance, panpsychism, hylomorphism and panpsychistic identism, (Presocratics, Giordano Bruno, Denis Diderot, Bernhard Rensch). Thus, we need another condition for a naturalistic theory of mind, e.g., a form of supervenience. Another reason for that is to exclude the possibility that two different mental states (tokens) are realized by or identical with one specific physical state (token).

The central intuition of supervenience is that "fixing" the physical fixes everything, or that nothing could have been otherwise without something physical having been otherwise. Supervenience holds that for every mental change there must occur a (simultaneous or preceding) physical change. Or, to be more precise, a set of mental properties supervenes on a set of physical properties if and only if any two persons that are indiscernible with respect to their physical properties are also indiscernible with respect to their mental properties, i.e., if any two persons that differ with respect to some mental property also differ with respect to at least one physical property. This is still a crude definition because one must clarify the scope of supervenience (cf. Beckermann, 1992): Does it hold in every possible world or even between possible worlds? Is it global or local? If externalism is true we cannot restrict the supervenience of mental events to neuronal (or bodily) events only, because identical neural events could have different contents in different environments (cf. II.4.). Thus, local supervenience seems to be too narrow and strong. On the other hand, global supervenience is much too broad and weak, because it does not seem plausible that every physical property within the scope of a person's relativistic light cone is relevant for the mental properties of that person. I do not assume for example that my mental states supervene on the properties of the thunderstorms in Jupiter's atmosphere or an interaction between two hydrogen atoms in the Large Magellanic Cloud at this very moment. Of course, this weak or global form of supervenience does not imply such crazy possibilities, but this example indicates that something has to be added or specified. Furthermore, supervenience alone is not sufficient for naturalism and hence a naturalistic theory of mind, because it is also compatible with epiphenomenalism, parallelism and occasionalism. Thus, we need an identity theory for a naturalistic theory of mind to exclude such a dualism, or at least a principle of physical exhaustion. However, as John Haugeland (1984, p. 119) has argued, it might be enough to get rid of "scientifically unmotivated, magically undetectable, and thoroughly bizarre" hypotheses by shifting the burdens of proof to the proponents of those hypotheses and accepting the heuristic rule "Don't get weird beyond necessity.

Nevertheless, fascinating questions and problems remain. For example, are our neural correlates of consciousness advanced enough to cope with a naturalistic world view (cf. Vaas, 1995b, 1996), and are they complex enough to understand their own complexity?

Acknowledgements: I am very grateful to Gary Schouborg and André Spiegel for valuable comments and Thomas Zoglauer for discussion. – This paper is based on a lecture given on June 20th 1998 in Bremen (Germany) at the international conference "Neural Correlates of Consciousness. Empirical and Conceptual Questions", organized by the Association for the Scientific Study of Consciousness.

References