THE NEURONAL BASIS OF CONSCIOUS EXPERIENCES IN THE ANIMAL KINGDOM

W. van de Grind

Comparative Physiology & Helmholtz Institute, Utrecht University,
Padualaan 8, 3584 CH Utrecht, The Netherlands.
e-mail: W.A.vandeGrind@bio.uu.nl

ABSTRACT

The study of conscious experiences in animals has been suppressed for a long time by the dominance of behaviorism and by the conviction that it would be impossible to study private experiences experimentally. Times have changed and viable techniques have been developed to use conscious experiences as experimental variables. I outline a conceptual model illustrating why conscious experiences have survival value and have been selected for. The model ties conscious experiences to learning and to potential messages for biocommunication. It is argued that ethics based on uneducated gut feelings is not a reliable guide to conduct and that we urgently need experimental studies of conscious experience, including studies on pain and suffering. Examples of promising approaches in the field of conscious visual experiences (percepts) are given to show that neuroscience has progressed far enough to start to solve puzzles regarding conscious experiences in the animal kingdom.

Keywords: Conscious experiences - Evolution of conscious experiences - Perception. - Pain and ethics

INTRODUCTION: CONSCIOUS EXPERIENCES AND THE NATURAL SCIENCES

Ever since Watson's initiation of the powerful movement of behaviorism (Watson, 1913) many scientists have stressed that we should study behaviour rather than consciousness or other postulated hidden internal processes. A recent defense of this position by Vanderwolf (1998) emphasises how mentalistic concepts have impeded scientific advance in the study of animal behaviour. This argument can be turned around, however, since the reign of behaviorism has stifled the scientific study of conscious experiences for more than half a century. Since our own conscious experiences are so undeniable, it seems unwise to dogmatically ban these interesting phenomena from scientific investigation. If "conscious experience" is an invalid concept in our theories of animal behaviour we should be able to prove so, rather than simply ban it *a priori* from our discourse.

I take it as a fact of (my) life that some of the activities in my nervous system lead to explicit personal experiences that I can voluntarily report, whereas most other activities in my brain are "inpenetrable", which means that I have no conscious access to them at all. Introspection is very limited and often misleading. We cannot introspect *why* we are in love or want to write a poem. Nevertheless, we know these things, so some basic form of introspection, of looking into oneself, exists. I can introspect my preferences in certain situations and many decisions pop up, but I cannot introspectively analyse the causes and reasons for this. It is perhaps amazing that we experience so little of what goes on in our brain, but it is too extreme to deny that we *have* inner experiences. An inner experience will be called conscious if it can be voluntarily and reproducibly reported or signalled in some form or other. Now the sceptic might retort that if my neighbour reports voluntarily that he has a headache, he could be a zombie who is programmed to occasionally say that, rather than a conscious person who experiences a headache. Yes, and I could be surrounded by hundreds of invisible men, but in fact I don't believe in zombies, nor in invisible men. In stead I use the analogy postulate in my social life.

The "analogy postulate" states that similar behaviour under similar circumstances is probably based on (or accompanied by) similar motivations and experiences. This postulate can be notoriously misleading if used loosely, without extensive analysis of what is meant by "similar" in each case. Despite these problems, there is no immediately obvious alternative and without the "analogy postulate" I would have no reason to assume that my neighbour reasons, loves, hates, has emotions and is a conscious person. The postulate is superior to the stance of the extreme sceptic, because it leads to falsifiable predictions about my neighbour's behaviour in future situations and after sufficient observations to a theory of my neighbour's psychological make-up. In brief it can lead to theories and models that really work, the hallmark of rationality and science. Of course the postulate can also be misused and lead to nonsense statements. (E.g. "Insects fly towards the light because they are curious"). Only after thorough study of the natural behaviour of the two species to be compared can one cautiously use the analogy postulate to generate hypotheses (rather than certainty) about the probable similarities and differences of inner experiences. This is the only direct bridge we have between inner experiences in different species, but if used with care it is at least a start. The next step can be to correlate the presumed inner experiences with measured brain activity in each of the species and analyse the similarities and differences against the background of knowledge about homologies in brain structure and function. Tedious, but solid empirical science, which is to be preferred to free speculations or dogmatic prejudices.

We share basic needs in life like eating, drinking, sex, health and safety, with the rest of the animal kingdom and none of these needs is found in computers. Thus the popular thesis that computers are (or will eventually be) more like humans than animals are, has always seemed preposterous to me. Computers don't even have a body! The idea seems to stem from religion rather than science. If god made man in his image and man makes a machine in his image, then surely the machine must have some of the godly properties as well. Consciousness, especially in connection with language, has always ranked high among the godly properties of humans that animals presumably lack and computers might eventually develop. I have no patience with these non-falsifiable theses and start from the usual biological perspective that humans evolved from and evolved like other animals, that the basic components of the nervous system (nerve cells, glia cells, synapses, transmitters, neuropeptides, etc) have evolved more than 300 million years ago and are universally present in the animal kingdom. Species differences stem from different numbers of nerve cells, different neuronal circuits, differences in niches, in bodily and behavioral adaptations. The "needs" or "drives" are the engines of survival so it should not surprise us that we share them with other animals. Without the urge to eat, drink, love, stay alive and well, none of the animal species would survive very long. These needs or drives form the universals of the brains of virtually all animals.

Searle (1984, 1992), has admirably argued against the popular but unfortunate idea that our nervous system is a kind of computer. More recently he (Searle, 1998) has attempted to correct a number of misunderstandings that might unjustly keep us from studying consciousness scientifically. I agree completely, but will not repeat his arguments here. Of the popular misunderstandings I think the most damaging is the idea that consciousness is an epiphenomenon rather than part of a causal process, and thus that it could not have played a role in evolution. The history of this misunderstanding is well-known. It started in 1870 with Shadworth Hodgson, a Darwinian who argued that conscious experiences are caused by the brain, but cannot in turn influence the brain processes. T.H.Huxley coined the name epiphenomenalism for this point of view and in a talk in 1874 entitled "The human as a conscious automaton" gave many analogies to drive the message home. He compared consciousness with the ticking of a clock, or the whistle of a steam engine, phenomena that are a consequence of mechanical processes but don't in their turn influence these processes. Interestingly it was a psychologist, the famous William James, who tried to ensure an evolutionary role for conscious experiences in his 1879-paper entitled "Are we automata?"

James gave the example of a child burning its hand at a stove. The hand will be withdrawn in a reflex, so there is no conscious experience before it is too late for the occasion. Pain comes too late, so why does it come at all? James points to the role of conscious pain as a negative reinforcer, as a signal to interrupt all other actions and direct all attention to the dangerous situation at hand. The pain interrupt-signal means: "Attention! Remember the sequence of events leading to the pain and prevent this in the future". Pain and other conscious experiences have survival value since they can mobilise the resources of the nervous system to learn what to do or not to do in similar situations in the future. If conscious experiences have survival value they must have been selected for in the process of evolution. The example by James brings the study of conscious experiences back into the realm of legitimate biological interests, and leads to the expectation that consciousness (and pain perception) must be wider spread than only in humans or only in primates. It is an empirical matter whether any given species has conscious experiences, pain percepts, the capacity to suffer. Pain has no doubt started as "pre-pain", as a warning signal to retreat, to minimise the risk of damage (see Broom, 2000, this issue). First in sufficiently complex modular nervous systems, with many parallel processes all doing their own job, does the need arise for a real massive interrupt to all modules, so that all or most resources can be used to learn about the causes of the problem and to prevent it in future. That is also where conscious experiences come in and where nociception becomes pain (Broom, op.cit.). In this connection I will suggest in the analysis-section (under 2) that conscious experiences are characteristic of the actions of the brain's operating system.

The study of consciousness has recently become more respectable in neuroscience and the number of papers on the possible neural mechanisms of conscious action and perception is increasing rapidly. Most of this modern work simply neglects the musings of philosophers and I will follow this bad habit, not because of lack of respect for philosophy or philosophers, but because of a lack of space. The neuroscientific approach is rather pragmatic and straightforward (you might want to call it "naive") and thus easier to do justice in a brief report, than the multifarious approaches of thinkers of various persuasions. Nevertheless, we need to spend some effort in analysing the concepts consciousness and conscious experiences to prevent misunderstandings and vagueness. That will be the first topic of the analysis section.

ANALYSIS

Types of consciousness and conscious experiences. Steps towards definitions

I will assume that the general state of activity of the brain can always be characterised by physiological means as being somewhere on the continuum from wakefuleness (W) to deep sleep or coma (C) and I will mockingly call this the WC-dimension (Van de Grind, 1997). I take it as axiomatic that the state of the brain needs to be closer to the W than to the C side on this scale to have conscious experiences and below I will take such a state of wakefulness for granted. Consciousness along this dimension (WC-consciousness) does not pose particular conceptual problems and it can be handled well in the operating theatre. It depends on the activity of the reticular activation system. It is a pity that the word "conscious" is used at all for this general and variable brain-state, since it is more like a precondition. It is fortunately also often called "arousal". Given that an organism is sufficiently aroused it might have "inner experiences", called conscious x or y, that accompany its measurable brain-body activity.

Delacour (1997) has attempted to summarise all aspects of the concept "consciousness" as used in neurobiology-related literature. Unfortunately this does not lead to the kind of conceptual simplicity that is so essential in experimental approaches. The term "consciousness" has so many connotations that it is almost useless in experimental studies. I will therefore preferably use the term "conscious experiences". They are the voluntarily and reproducibly reportable (communicable) experiential items that appear against the background "stream of consciousness" (arousal). The discrete conscious experiences can be raw feelings (qualia, emotions) or have specific content (objects, thoughts, plans), two classes termed phenomenal (P-) and access (A-) conscious experiences, respectively, by Block (1995). Even though I like this classification I will simply call any voluntarily reportable inner experience a conscious experience. Neuroscience can, I think, study some or most of these concrete items, like the quale "red" or "motion to the left" or the conscious experience of an object or a spatial relation or a decision to act.

Much has been made of self-consciousness in the philosophical literature and many advertise this as the highest form of consciousness. I am sceptical of such an exalted qualification, since all animals have extensive self-knowledge and it is hard to see why self-knowledge should be so different from knowledge of the external world. Without self-knowledge the lion would eat its own front paws while enjoying its bloody prey and small animals would unhesitatingly start fights with much larger rivals. As with knowledge of external affairs most self-knowledge will be inpenetrable, so self-knowledge does not prove self-consciousness. However, the point is that both types of knowledge are essential to survival and to communication and they can therefore be expected to have co-evolved. If items from one knowledge domain can lead to conscious experiences I see no reason why items from the other domain should not. Probably pain and the capacity to suffer, forms of self-knowledge, are more urgent to survival than most forms of object-knowledge. Thus conscious self-experience might even be a more basic (evolutionary older) instead of a higher (more recent) form of conscious experience. The Gallup mirror test (Gallup, 1970, Gallup *et al.*, 1995, criticised by Heyes, 1994, 1995) is interesting to test the capacity of animals to handle difficult dynamic visuo-motor mirror transformations, but I cannot see why it should say much about self-consciousness in animals. Also certain patients with brain damage do not recognise themselves in mirrors, but are nevertheless self-conscious.

Language, a human specialisation, can of course also influence conscious experiences. Take the following statement: I am conscious of the fact that my neighbour is conscious of the fact that I am conscious of myself and of the fact that the neigbour is conscious of my being conscious. I don't think we could think such a thought without language. Language certainly expands both our potential knowledge-base, our possibilities of reasoning and of voluntarily reporting inner feelings. However, language and conscious experiences can be doubly dissociated (Milner and Rugg, 1992). This means that language is neither a precondition nor a diagnostic of (all forms of) conscious experience, but it might bring some additional flavour to it all. Here I will neglect the influence of language in an attempt to approach the topic of conscious experiences in a species-independent fashion.

A possible function for conscious experiences

A thesis proposed by a great zoologists of the previous century, J.Z.Young, can serve as my starting point. He proposed that brains contain plans for action, called "brain programs" (Young, 1978), which regulate the life of the brain-owners and specifically help them to survive. One should not construe this as another misguided analogy between computers and brains. Computer algorithms form but one specific narrow class of programs, but there are many more types of program. A causal structure like a frame and wheels, pedals and a chain, together called a bicycle, can be viewed as a "program for locomotion". Social events are usually based on some program and a neuronal circuit stabilising our body-temperature in environments from the tropics to the poles is a program. A program is an ordered set of conditional action commands or actions, measurement-commands or measurements and situation-dependent decisions, together leading to some specific global effect. A program has a goal and this is the characteristic difference between a program and your average physical process, like meltwater streaming down a mountain-slope. I like the brain program metaphor, because it emphasises the purposiveness of brain circuits and the conditionality of their actions.

Modern neuroscience has made it abundantly clear that the brain consists of many modules working in parallel and each consisting of a large collection of brain programs that do their things when needed. We have brain programs to search specific visual information in complex scenes by following certain eye-movement strategies, we have brain programs for catching balls or chewing food, for stabilising glucose-levels and temperature, for placing our feet while balancing on a fallen tree to cross a river and for gaining access to the other sex. Many, if not most, of these brain programs do their glorious work completely automatically, without allowing the owner of the brain any introspective access to their inner workings. Zeki and Bartels (1999) deduced a number of propositions on the functional architecture of the visual system from an extensive review of work they and others did to clarify the mechanisms of conscious visual perception. They characterise the visual brain as a parallel, distributed system with functionally specialised modules (they call them nodes) that are fairly autonomous. (Their propositions 1 and 15 combined). This is not expected to be different for other brain systems. I cannot here review all the propositions by Zeki & Bartels, but I can strongly recommend reading their paper. Their central thesis is that all processing nodes have their own microconsciousness, but they admit not having proven this. Like Crick and Koch (1998) I tend to think more in terms of a central office, in my terms an "operating system" that is responsible for the conscious experiences. It is reminiscent of the "central scanner" proposed quite some time ago by Armstrong (1968). Yet, the proposal by Zeki & Bartels is a viable alternative and one can even imagine both proposals to reflect part of the solution.

There have been heated debates in the past about the idea of some central consciousness module in the brain. The proposal has often been viewed as a modern incarnation of Descartes' proposal of a nonmaterial mind reading and commanding the brain through the pineal gland. Dennett and Kinsbourne (1992, see also Dennett, 1991), with feeling for drama, characterise it as the "Cartesian Theatre". All this excitement is a bit peculiar, because from an engineering point of view one can often choose between a distributed or concentrated location of network components without changing the operating principles, the network structure. My proposal, to be worked out below, is to assume that conscious experiences are characteristic of activity in the brain's operating system. This proposal is in principle neutral as to the question of a distributed or concentrated localisation of the BOS (Brain Operating System). However, it would really be nice if the localisation were not too messy since a strongly scattered system would make it harder to study the question in what sense BOS-neurons are different from the other nerve cells of the brain. Surely, they must be, since their activity supports our experiences, whereas all other nerve cells appear to lack this ability. For the time being I hope the BOS is concentrated in one place, e.g. somewhere in the prefrontal cortex and underlying structures. The BOS is certainly nothing like Descartes' conscious mind. It only synchronizes and coordinates and like any operating system is far from omniscient. It needs to detect and prevent simultaneous conflicting commands (like "sit down" and "run"). It gets strategic advice from modules that the BOS cannot look into, so it knows the advice (our intuition), but not the reasons for the advice.

Psychology studies serial processing in the brain, like attention, voluntary action, conscious experiences, and neuroscience studies parallel processing (see Baars, 1997, 1998). This fits in well with the core of my proposal, where the BOS is the bridge to psychology, the bottleneck process. If a massively parallel system is to work well one needs much (but not total) autonomy of the component systems, but also a well-tuned bottleneck, a serial system, for important decisions. Building a house is a good example of a parallel process in which all the specialised firms do their job almost autonomously, save for certain triggers (timing) from the contracter. Only the contracter knows the global state of affairs but he need not know the details of the work of plumbers and carpenters, who in their turn don't bother to tell him what pipes or nails they use to get their job done. The BOS is uninformed about the details, not able to check how the various specialists do their jobs, but capable of starting programs or changing them, giving interrupts, reading status reports (presumably as condensed as oneliners), etcetera. Superficially my proposal (of an operating system) looks similar to the one by Johnson-Laird (1993), but I don't accept the idea of a strong analogy between computers and brains. "Operating system" is to me a general term that, like the term "program", is useful for a far broader range of systems than only computers. It differs from "board of directors" or "central executive" in that it does not suggest a human of group or humans (that is very complex systems) doing the coordinating job, but a mechanical structure that in itself follows simple rules of behaviour rather than being homunculus-like.

Brains can learn and learning in the context of the brain-program and BOS metaphors means either tuning an existing brain program parametrically, or constructing a new brain program from scratch, or a bit of both: adapting and combining brain programs. If programs in your parietal and temporal cortex could be constructed independently of each other, each with some local goal structure, chaos might result. Some supervision is required. Moreover, with many parallel modules there might be a risk in the case of an approaching predator to start programs for climbing a tree and running away simultaneously, probably resulting in the untimely death of the organism. Therefore a BOS is needed to keep the various parallel modules and their programs in check, to coordinate if necessary. That is in my view where consciousness enters the scene. Conscious experiences are activities in the BOS. Since there is no external programmer, the BOS must be self-(re)organising, self-checking, and most specialised programs must be self-organising within their own modules (brain regions), perhaps with a little help from the BOS. In fact the BOS-metaphor can be taken a step further in the case of social animals, where external signs of inner feelings can be called the "desktop" functions of the BOS. They are made available to inform external observers about the status of the BOS.

Why should BOS-activity correlate with inner experiences? I think the main point here is the possibility of generating voluntary reports, of communication. The BOS also embodies our link with other organisms. Our displayed anger and threat can chase them away, our friendly smile can attract them. Biocommunication is in this model intimately tied to obvious BOS-functions of promoting survival by dedicating certain resources to certain tasks. This is the package-deal. If you have a BOS, and you can communicate, it has survival value to make the BOS-activity available for voluntary communication. Conscious experiences are potential sources of communication. Donald Griffin (1984) who was one of the first to urge that animal consciousness should be taken seriously again by biologists and studied experimentally, emphasised that one should "listen in" on communication between animals to learn more about their conscious experiences. My BOS-model gives a rationale for this idea in that it ties conscious (voluntarily reportable) experiences to biocommunication.

It might seem strange to assume that BOS-activity can cause conscious experiences and present opportunities for communication, while we know how difficult it is to communicate conscious experiences in language. However, I think that this is because BOS-activity does not generate language, but only messages in the form of raw feels, action decisions or percepts. The language system has to "translate" these proto-messages as good as it can, and it often produces suboptimal results. As Roland (1994) remarks, common language is poor, it does not even have a specific vacabulary for smells. Thus it appears that the bottleneck in communicating about our conscious experiences is their translation into language. Blushing or laughing is much more direct and is universally understood by members of our own species. In fact this is precisely the idea put forward by Ramachandran and Hirstein (1997) who think that the epistemological barrier between conscious experiences of somebody else and me is merely the necessity to translate from her neuron firings to her spoken language to my auditory system's neuron firing. From my auditory system the activity will never reach the BOS-module in my brain that corresponds to the module originating the message in the sender's brain. If we could use a direct neuron cable between corresponding regions of the two brains I would, according to these authors, immediately exclaim "oh my God, I see what you mean". This is their solution to the so-called hard problem of consciousness, the problem that we cannot communicate to others what a conscious experience really feels like (Nagel, 1974). Of course things are never as simple as they seem, so one could certainly be sceptical about this, but the idea has a certain plausibility.

Why and how a general theory of conscious experiences is indispensible

Roland (1994) has criticised the use of metaphors drawn from business, electronic engineering and computer sciences in neuroscience, because, as he maintains, they do not oblige the author to take the ontological properties of the nervous system into account. He wants models to be "ontologically committed at many levels to the system being modelled" (p.25, op. cit.). I think this is demanding too much and I don't know a single example of successful modelling in the natural sciences that conforms to this type of dogma. In neuroscience, neuron-membrane models are always formulated in electric engineering metaphors. They consist of batteries, condensors and resistors. Such models allow a high precision of quantitative predictions of membrane behaviour and are indispensible. The Hodgkin Huxley equations are not framed in neuroscientific terms but in mathematical language, spike propagation along axons is described in terms of signal propagation along leaky cables, the same formalisms as used in electrical engineering for transatlantic cables. Neuroscience as we know it today is simply unthinkable without the metaphors, analogies, and formal descriptions stemming from electrical engineering! Thus I think Roland is simply mistaken. Neuroscience does not differ much from electrical engineering. Neuroscience has no special ontological status and the nervous system obeys the same laws of electricity and electrochemistry that are used in engineering. The engineering experience in designing complex circuits can thus be of great help in designing models of neuronal circuits. All this is, of course, not to deny that neuroscience needs additional concepts as well, such as those from biochemistry and molecular biology (neuropeptides, proteins, genes, and what not).

Similarly, if one wants to think about the purposive cooperation of the many almost autonomous modules in the brain one needs metaphors as tools of thought (again: not as articles of faith). What do we know about organising complex systems? Well, large businesses and computers are organised complex systems. The digital serial computer itself is not in any sense a useful or acceptable model of the brain. Roland (1994), Searle (1984, 1992, 1998) and many modern neuroscientists (including this author) can easily agree on this. However, computer science also includes studies on how to organise large sets of semi-autonomous agents. It has successfully taken its metaphors (e.g. "program", "memory", "interrupt", "word") from daily life and from neuroscience! In addition, there are many common interests ahead (self-programming systems, self-organising operating systems, operating system design for parallel systems, etcetera). Metaphors and analogies are the trademark of science, from strings and black holes to the brain operating system (BOS) discussed above. A model is not a simplified description of the system under study (Roland, p.25, op. cit.), but a tool of thought allowing us to predict the outcome of innovative experiments on the system in order to gain new insights!

Baddeley and Hitch (1974) proposed a very influential model of "working memory" consisting of a central executive (an attentional control system) with two slave systems, a phonological loop and a visuospatial sketchpad (see also Baddeley, 1998; and for criticism Roland, 1994). Neglecting the phonological loop, which is only relevant for humans, the proposal is simply that there is an attentional control system and that visuospatial information can be held available ("on line") to this control system while choices are being made for redirecting attention or while reasoning is going on. I include such functions in my BOS, of course. The point here is that this simple proposal was very fruitful in designing experiments that led to new insights and data. A kind of visuospatial sketchpad or working memory system has since been found in the prefrontal cortex of primates, a.o. by Goldman-Rakic (1988, 1992, 1995). She found cells in Brodman's area 46 with memory fields. These cells become and stay active if and so long as their specific position in the visual field needs to be kept in working memory. Also cells were found with object rather than position memory (Wilson *et al.*, 1993). These and similar findings have strongly stimulated interest and research in the prefrontal cortex of awake task-performing monkeys. Such work allows the direct correlation of voluntary reports by trained animals and single cell activity in the central executive or working memory or BOS system. At present this seems a viable approach to the detailed neurobiology of conscious experiences.

Interestingly, a kind of consensus is developing that Baddeley's central executive needs to be further subdivided (Roberts *et al.*, 1996). That is exactly the idea of my BOS, which needs to have many control modules (not only for directing attention), kinds of working memory (not only visuospatial), and possibilities for routing interrupts (e.g. when pain occurs), command signals (e.g. when a "voluntary" action needs to start), and request signals (e.g. if information from long term memory stores needs to be retrieved). It is hard at present to formulate all properties that must be inherent in the BOS, but one boundary condition is a priori clear: it should not become an omniscient homunculus-like system, but a coordinative structure working with limited information and providing limited instructions. Autonomy of the expert modules should be as high as possible without breaking the overall coherence that is essential to survival of the organism. Thinking through what kind of interrupt signals, command signals, request signals are necessary to perform certain tasks and what kind of information is minimally needed from specialised modules (which are localised and characterised at a high pace in modern neuroscience), one can design experiments to test the resulting hypotheses. This will slowly but surely lead to a characterisation of the functional and actual architecture of the BOS.

I think one needs this (otherwise minimally restrictive) metaphor of a BOS-system to sharpen the boundary conditions of experiments, to develop expectations about the cooperation of "central" functions like attention control, working memory, long-term memory, emotion and communication. It also immediately illustrates why it is reasonable to expect a gradual evolution of the number and variety of conscious experiences. Animals with few neurons, or with, say, ladder-like nervous systems, might lack the necessity to have a BOS and they might hardly communicate. Thus they might not need and have conscious experiences. The BOS-idea and the requirement of voluntary reportability might give us a handle on presumed inner experiences in other animals. It would make it a truly empirical question what kind of conscious experiences a given species might have, how important they are (number of neurons dedicated to it or size of the axons carying interrupt signals and the like). Of course the criterion itself (voluntary and reproductible reportability) should remain questionable as well. It is a kind of best bet based on the analogy postulate, nothing more, but certainly nothing less. Pain, as a major potential interrupt and "remember this" command and conscious (voluntarily reportable) experience, that is communicated in some species (distress signals), but not in others (Broom, 2000), must be a strong contributor to BOS-activity.

Laws and rules against inflicting pain in animal experiments are so severe that we are probably forever stuck with articles of faith and ethics that have no empirical basis. Ryder (2000) for example states "*Pain is pain regardless of the species of the sufferer. (Pain in this context, covers all forms of suffering including fear, boredom and distress.) Try to act, therefore, as though human pains count for no more than nonhuman pains*" I do not a priori count human pains for more than nonhuman pains. Yet, I would like to see scientific evidence that, say, an earthworm is capable of suffering pain, before equating the cutting in two of a living earthworm with the cutting in two of a living human. Moreover, fear, boredom and distress are natural phenomena and I have suffered them a lot in my life. Yet I see no reason to forbid my collegues, family, government to sometimes bore me for a while or even cause fear or distress in me. The case of animal ethics is (in present times in western societies, most notably England) strongly overstated and if we don't try to come to our senses (!), the time will come that farmers are arrested because their chickens and cows are bored and don't get a proper education. Animal experiments are the only source of concrete knowledge on conscious experiences in animals and thus on pain and suffering in animals. I therefore think animal experiments are indispensible for the good of animals. (I am not so much in favour of animal experiments for the good of humans, although I do not reject them categorically). One can minimise the risk of suffering in animal experiments by first studying other potential conscious experiences than pain, but eventually some form of pain stimulation might also be needed. This could be limited to dentist-like experiences or to consequences of natural suffering (e.g. after fights in nature). However it is done, scientific insight in animal pain and animal suffering is sorely needed, lest we are stuck with the truly silly consequences sketched above of a priori dogmas (ethical rules) based on uneducated gut feelings.

Useful examples of the neuroscientific study of conscious experiences

I will restrict this discussion to two promising examples of the study of conscious experiences in nonhuman primates. One can easily envisage extensions of these approaches to other species, albeit certainly not all the way back to flat worms. Both examples are about vision, but it is not difficult to see how they could be extended to studies of other perceptual systems. Vision is evolutionary old and found in all animal phyla, unlike for example hearing, which is only found in vertebrates and arthropods. Vision therefore has a certain advantage if the possibility of conscious experiences (such as visual percepts) needs to be probed in various lineages.

First I will paraphrase work of Newsome and co-workers (e.g. Salzman *et al.*, 1990; see also the review by Parker and Newsome, 1998). Visual brain area MT or V5 (visual cortex area 5) is filled with cells reacting optimally to motion. If a collection of randomly positioned dots drifts coherently in some direction on a screen, cells tuned to that direction of drift react optimally. Cells tuned to the opposite direction are suppressed or show an undisturbed spontaneous activity and cells tuned to directions in between show responses in between. Now if the amount of coherence in the drifting dots is decreased by randomly refreshing a certain percentage of dots on every frame of the movie, the responses decrease gradually with decreasing coherence. On the whole these cells can explain the psychophysically measured responses of the trained monkey or a human to such stimuli very well.

Newsome and his coworkers then introduced electrical microstimulation, to influence the behaviour of groups of cells around the stimulation electrode. Since cells tuned to certain motion directions appear to cluster in V5 this allows one to suppress virtually all cells tuned to a certain direction, or with the opposite current-direction to stimulate them all. If the monkey detects the motion direction well at a certain level of coherence, as measured psychophysically from its bar-pressing behaviour, inhibitory electrical stimulation to cells tuned to the stimulus direction decreases the detection performance of the monkey. It is possible to make the inhibition strong enough to get guessing behaviour of the monkey. Thus, if the considered cell groups are active the monkey reports the perception of motion in the proper direction, if they are silenced by electrical suppression the monkey reports that it is guessing, not perceiving, the motion direction. The cells are necessary for perception of their corresponding motion direction! Next these cells were excited electrically, in the absence of their visual preference stimulus on the screen, and the monkey proved to report the presence of that visual motion direction. The cells are therefore also sufficient for the corresponding percept.

Of course experiments like these need many controls and quantitative analysis, which have indeed been carried out, but this is not the place for a detailed review of the experiments. The point of importance here is that certain specific cell groups in V5 are both necessary and sufficient for a conscious (voluntarily and reproducibly reportable) motion percept, that corresponds to the stimulus preference of the cells under study. Only after extensive quantitative studies of brain areas like V5 and others in awake performing monkeys does it become possible to reliably correlate conscious perception and cell activity directly. Developments in the last decade have shown that this approach really works quite well. It can be extended in principle to other species, provided they can be trained to perform in the required psychophysical test.

A second example involves binocular rivalry. Imagine first a mirror construction enabling the independent (so-called "dichoptic") stimulation of each of the two eyes. Then suppose one eye sees vertical stripes and the other horizontal stripes. The perceptual result (if you have normal binocular vision) will be binocular rivalry. Sometimes you see vertical stripes, then again horizontal stripes and sometimes you see a kind of piecemeal mixture of vertical and horizontal patches. For relatively small stimuli you only see an alternation of vertical and horizontal stripes, which means that sometimes one eye dominates and then again the other. If observers press a button to indicate which percept dominates at the time one can study the distribution of dominance periods of the two eyes. It proves that this is a gamma distribution and that one can not influence the distribution by "will power", that is by trying to force oneself to more often see, say, vertical stripes. The same experiment can be done in awake performing monkeys while measuring responses of cells in various visual areas to the striped patterns. One finds the same gamma distribution of eye-dominance periods. In most visual areas cells have a preference for either vertical or horizontal stripes (or other orientations, but that is not relevant to this story).

Schall and Logothetis, but especially Logothetis (e.g. Logothetis, 1998a, b) used dichoptic stimulation to study responses of cells in various visual areas of awake performing monkeys to such stripe (and other) patterns, in dependence of what the monkey reported to be seeing at the time. Leaving out all subtilities and experimental controls one can summarise the results as showing that cells in V1 (the primary visual cortex) mostly follow the outside world, cells in temporal visual areas mostly follow what the monkey reports seeing. That is, if a temporal cortex visual cell responds strongly to vertical gratings it is only active during the periods that the monkey reports seeing vertical stripes. A V1-cell with the same stimulus preference, however, will simply signal the presence of such a grating regardless of the monkey's percept. Again we see a good opportunity to study the relation between activity in the nervous system and conscious experiences. There are more examples in the modern literature, some even more striking, but as yet often less solidly embedded in huge amounts of independent controls. For the message I want to convey here, the above probably suffices. It has become possible to directly correlate conscious experiences with activities in individual brain cells and small groups. We can expect exciting results in this area of research in the years to come and they will make consciousness research a fascinating experimental discipline in neuroscience.

Functional-MRI-studies are of course also contributing strongly to the study of conscious experiences. One can for example try to localise the area(s) of the brain responsible for flips of dominance in binocular rivalry, or areas that are most active while subjects are doing calculations or solving a puzzle or having a tooth-ache. The resolution in time and space of these studies is relatively low, however, in comparison to single-unit recording studies. Moreover, it is still virtually impossible (despite the invention of transcranial electrical stimulation) to influence neuronal activity during fMRI, like one can during micro-electrode studies with electrical microstimulation (or pharmacological stimulation). Moreover, fMRI cannot look deep into the brain, whereas microelectrodes can. Thus I think that neuroscientific consciousness research needs single unit recording studies in awake performing animals, for decades to come, if not forever. It is an urgent mission, also from the viewpoint of animal welfare, to scientifically explore conscious experiences in nonhuman animals. I hope that scientific institutions and animal welfare groups will have the good sense to support such research, rather than condemn us to unscientific prejudices and schools for higher education of ducks, worms and flies.

CONCLUSIONS

1. Conscious experiences are voluntarily and reproducibly reportable, in contradistinction with nonconscious experiences, which are introspectively inpenetrable.
2. Voluntarily reportable experiences can be used as variables in experiments and examples have been given.
3. Models, like my BOS-model, are necessary to structure the field of study of conscious experiences across species and through evolutionary time.
4. Experimental insights and testable models should replace the uneducated gut feelings that now dominate ethics and discussions regarding animal welfare.

REFERENCES

Armstrong D.M. (1968). *A materialist theory of the mind.* Routledge & Kegan, London.

Baars B.J. (1997). Some essential differences between consciousness and attention, perception and working memory. *Consciousness & Cognition 6*, 363-371.

Baars B.J. (1998). Metaphors of consciousness and attention in the brain. *Trends in Neuroscience 21 (2),* 58-62.

Baddeley A.D. (1998). Recent developments in working memory. *Current Opinion in Neurobiology 8*, 234-238.

Baddeley A.D., Hitch G (1974). Working memory. In: G.A. Bower (Ed.). *The psychology of learning and motivation*, Academic Press, New York, 47-89.

Block N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences 18*, 227-287.

Broom D.M. (2000). The evolution of pain. *Vlaams Diergeneeskundig Tijdschrift (Flemish Veterinary Journal), 69,* 1-5.

Crick F., Koch C. (1998). Consciousness and neuroscience. *Cerebral Cortex 8 (2)*, 97-107.

Delacour J. (1997). Neurobiology of consciousness: An overview. *Behavioural Brain Research 85*, 127-141.

Dennett D.C. (1991). *Consciousness explained.* Little Brown, Boston.

Dennett D.C., Kinsbourne, M. (1992). Time and the observer - the when and where of consciousness in the brain. *Behavioral and Brain Sciences, 15*, 183-247.

Gallup G.G. (1970). Chimpanzees: self-recognition. *Science 167*, 86-87.

Gallup G.G., Povinelly, D.J., Suarez, S.D., Anderson, J.R., Lethmate, J., Menzel, E.W. (1995). Further reflections on self-recognition in primates. *Animal Behaviour 50*, 1525-1532.

Goldman-Rakic P.S. (1988). Topography of cognition: parallel distributed networks in primate association cortex. *Annual Review of Neuroscience 11*, 137-156.

Goldman-Rakic P.S. (1992). Working memory and the mind. *Scientific American 267 (3)*, 73-79.

Goldman-Rakic P.S. (1995). Cellular basis of working memory. *Neuron 14*, 477-485.

Griffin D.R. (1984). *Animal Thinking.* Harvard University Press, Cambridge Mass.

Heyes C.M. (1994). Reflections on self-recognition in primates. *Animal Behaviour 47*, 909-919.

Heyes C.M. (1995). Self-recognition in primates: further reflections create a hall of mirrors. *Animal Behaviour 50*, 1533-1542.

Johnson-Laird P. (1993). *The computer and the mind.* Fontana Press, London.

Logothetis N. (1998). Single units and conscious vision. *Phil. Trans. R. Soc. Lond. B 353*,1801-1818.

Logothetis N. (1998). Object vision and visual awareness. *Current Opinion in Neurobiology 8*, 536-544.

Milner A.D., Rugg M.D. (1992). *The neuropsychology of consciousness.* Academic Press, New York.

Nagel T. (1974). What is it like to be a bat? *Philosophical Reviews 83*, 435-450.

Parker A.J., Newsome W.T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annual Review of Neuroscience 21*, 227-277.

Ramachandran V.S., Hirstein W. (1997). Three laws of qualia: what neurology tells us about the biological functions of consciousness. *Journal of Consciousness Studies 5 (5-6)*, 429-457.

Roberts A.C., Robbins T.W., Weiskrantz L. (Eds.) (1996). Executive and cognitive functions of the prefrontal cortex. *Philos. Transactions of the Royal Society of London B 351*, 1387-1527.

Roland P.E. (1994). Obstacles on the road towards a neuroscientific theory of mind. *Journal.of theoretical. Biology. 171*, 19-28.

Salzman C.D., Britten K.H., Newsome W.T. (1990). Cortical microstimulation influences perceptual judgments of motion direction. *Nature 346*, 174-177.

Searle J.R. (1984). *Minds, brains and science.* Harvard University Press, Cambridge Mass.

Searle J.R. (1992). *The rediscovery of the mind.* MIT Press, Cambridge Mass.

Searle J.R. (1998). How to study consciousness scientifically. *Philosophical Transactions of the Royal. Society of London B 353*, 1935-1942.

van de Grind W.A. (1997) *Natural intelligence* (in dutch). Uitgeverij Nieuwezijds, Amsterdam.

Vanderwolf C.H. (1998). Brain, Behavior, and Mind: What do we know and what can we know. *Neuroscience & Biobehavioral Reviews 22 (2)*, 125-142.

Watson J.B. (1913). Psychology as a behaviorist views it. *Psychological. Revue. 20*, 158-177.

Wilson F.A.W., Scalaidhe S.P.O., Goldman-Rakic P.S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science 260:* 1955-1958.

Young J.Z. (1978). *Programs of the brain.* Oxford University Press, Oxford UK.

Zeki S., Bartels A. (1999). Toward a theory of visual consciousness. *Consciousness and Cognition 8*, 225-259.