Consciousness Studies

From Wikibooks, the open-content textbooks collection

Edition 1.0 12th March 2006

Note: current version of this book can be found at
http://en.wikibooks.org/wiki/Consciousness_studies

Principle authors

RobinH, JWSchmidt & Shamantics (Wikibooks Users).

Acknowledgments

Many thanks to Dr Alex Green who has allowed the use of any work for which he holds the copyright. About 50% of the original book was taken or adapted from Green's material which got us off to a flying start. Dr. Angell O. de la Sierra contributed the section on the BPS Model. Thanks are also due to anonymous contributors.

Introduction

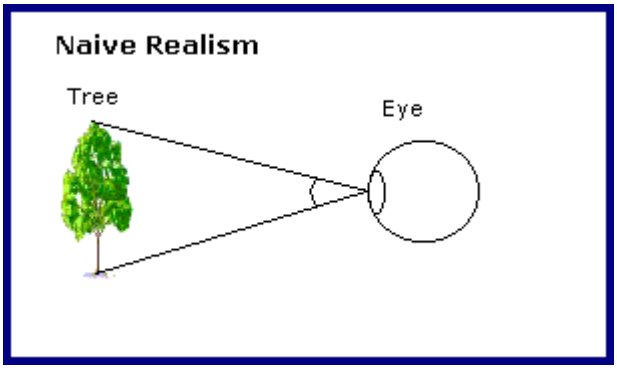
Everyone has their own view of the nature of consciousness based on their education and background. The intention of this book is to expand this view by providing an insight into the various ideas and beliefs on the subject as well as a review of current work in neuroscience. The neuroscientist should find the philosophical discussion interesting because this provides first-person insights into the nature of consciousness and also provides some subtle arguments about why consciousness is not a simple problem. The student of philosophy will find a useful introduction to the subject and information about neuroscience and physics that is difficult to acquire elsewhere.

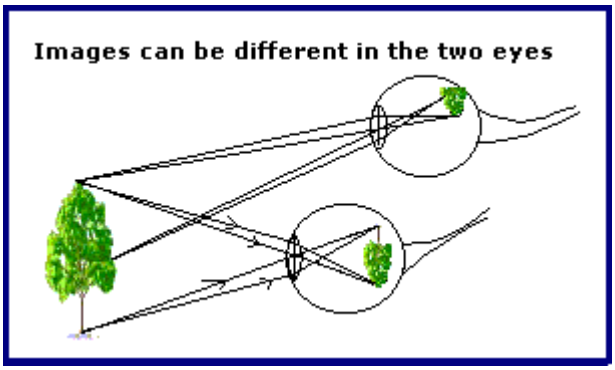
It is often said that consciousness cannot be defined. This is not true; philosophers have indeed defined it in its own terms. It has two principle components: firstly **phenomenal consciousness** which consists of our experience with things laid out in space and time, sensations, emotions, thoughts, etc., and secondly **access consciousness** which is the processes that act on the things in experience.

As will be seen in the following pages, the issue for the scientist and philosopher is to determine the location and form of the things in phenomenal consciousness. Is phenomenal consciousness directly things in the world beyond the body, is it brain activity based on things in the world and internal processes, "a sort of virtual reality", or is it some spiritual or other phenomenon?

A note on Naive Realism

Children tend to believe that the world is identical to the world that they see and feel. A very young child might even think that a curiously shaped shadow is a monster or be fooled into thinking that there really are people inside a television set. Older children with a smattering of geometry tend to believe that they have a 'point eye' that sees the world. Physical considerations show that ideas such as these are highly contentious; we have two eyes with different images in each, normally the only images in the world are created by optical instruments such as the eye and the photons that carry light to the observer cannot and do not all exist at a single point. Some of the discrepancies between the physical reality and our experience are shown in the illustrations.



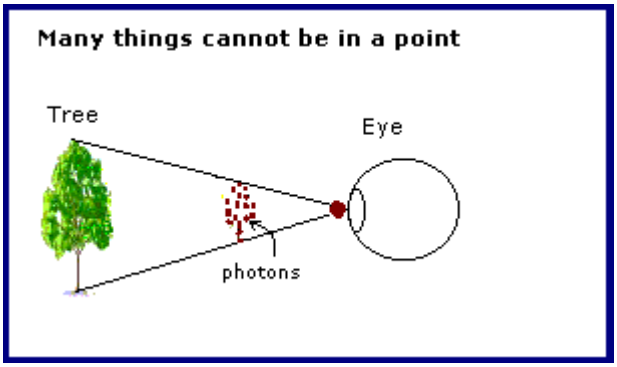


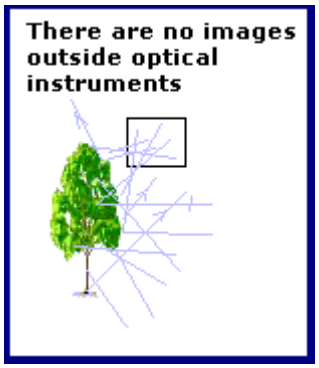
The naive realist idea of perception involves a point eye looking at a geometrical form. But the physics is different; there are two eyes with sometimes very different images in each.

Light falls all over the cornea from every part of a view - there is no 'point eye'.

.

The cloud of photons that compose light must get in the way of the view but naive realism neglects this, regarding the photons as somehow transparent yet gathering as an impossible group of millions of photons in a viewing point





Light rays go everywhere; it is only after light has passed through an optical instrument such as the eye that an image is formed. Hold up a sheet of paper - there are no images on it.

The illustrations show the nature of one of the most difficult problems studied by neuroscience: how can the images on the two retinas become experience? How can we imagine things or experience dreams and hallucinations?

A degree of Naive Realism is a sensible idea for coping with the everyday problems of working and living. Most physical scientists and people in general are, to some extent, Naive Realists until they study the biology of sensation and the problems of perception and consciousness. There is often a suspicion, or even fear, amongst Naive Realists that any analysis of conscious experience is a suggestion that the world does not exist or everything is imaginary. These fears are unfounded: Neuroscience is a study of the part of the physical world represented by brain activity and is part of medicine.

Other uses of the term "Consciousness"

This book is about a specific meaning of the word consciousness that is of interest to neuroscientists. This is defined carefully in the sections that follow. However, the term "consciousness" has been used by philosophers to describe the functional interaction of a person with the world, especially social and political interaction, and also to describe a person's moral conscience. The text below, based on the entry for "Consciousness" in Wikipedia, describes this other use of "consciousness".

"Consciousness" derives from the Latin word 'conscientia' which primarily means moral conscience. In the literal sense, "conscientia" means knowledge-with, that is, shared knowledge. The word first appears in Latin juridic texts by writers such as Cicero. Here, conscientia is the knowledge that a witness has of the deed of someone else. In Christian theology, conscience stands for the moral conscience in which our actions and intentions are registered and which is only fully known to god.

Locke's "forensic" notion of personal identity founded on an individual conscious subject would be criticized in the 19th century by Marx, Nietzsche and Freud from various different angles. Martin Heidegger's concept of the *Dasein* ("Being-there") would also be a tentative

to think beyond the conscious subject.

Marx considered that social relations ontologically preceded individual consciousness, and criticized the conception of a conscious subject as an Ideology as an instrument of an ideological conception on which liberalism political thought was founded. Marx in particular criticized the 1789 Declaration of the Rights of Man and of the Citizen, considering that the so-called individual natural rights were ideological fictions camouflaging social inequality in the attribution of those rights. Later, Louis Althusser would criticize the "bourgeois ideology of the subject" through the concept of *interpellation* ("Hey, you!").

Nietzsche, for his part, once wrote that "they give you free will only to later blame yourself", thus reversing the classical liberal conception of free will in a critical account of the genealogy of consciousness as the effect of guilt and "ressentiment", which he described in *On the Genealogy of Morals*. Hence, Nietzsche was the first one to see how much the modern notion of consciousness was indebted to the modern system of penality, which judged a man according to his "responsibility", that is by the consciousness through which acts can be attributed to an individual subject: "I did this! this is me!". Consciousness is thus related by Nietzsche to the classic philosopheme of recognition which, according to him, defines knowledge.

According to Pierre Klossowski (1969), Nietzsche considered consciousness to be a hypostatization of the body, composed of multiple forces ("Friedrich Nietzsche, The Will to Power"). According to him, the subject was only a "grammatical fiction": we believed in the existence of an individual subject, and therefore of a specific author of each act, insofar as we speak. Therefore, the conscious subject is dependent on the existence of language, a claim which would be generalized by critical discourse analysis (see for example Judith Butler).

Michel Foucault's analysis of the creation of the individual subject through disciplines, in *Discipline and Punish* (1975), would extend Nietzsche's genealogy of consciousness and personal identity - i.e. individualism - to the change in the juridico-penal system: the emergence of penology and the disciplinization of the individual subject through the creation of a penal system which judged not the acts as it alleged to, but the personal identity of the wrong-doer. In other words, Foucault maintained that, by judging not the acts (the crime), but the person behind those acts (the criminal), the modern penal system was not only following the philosophical definition of consciousness, once again demonstrating the imbrications between ideas and social institutions ("material ideology" as would call it Althusser); it was by itself creating the individual person, categorizing and dividing the masses into a category of poor but honest and law-abiding citizens and another category of "professionals criminals" or recidivists.

It can be seen that this other meaning of the term consciousness is connected to the idea of 'consciousness as conscience' in which conscious thoughts may or may not be used in the processes that might be called free will. It is not considered in this book.

Part I: Historical Ideas

This section is an academic review of major contributions to consciousness studies. Readers who are interested in the current philosophy of consciousness will find this in Part II and readers interested in the neuroscience of consciousness should refer to Part III.

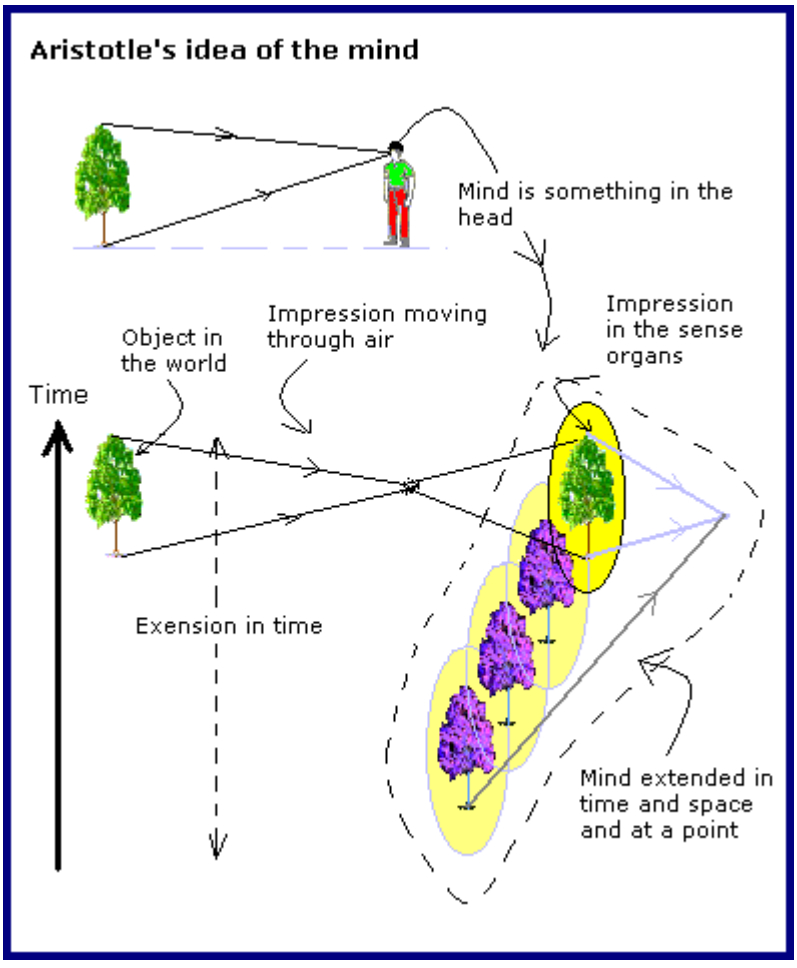
Aristotle. (c.350 BC). On the Soul.

(De Anima) http://psychclassics.yorku.ca/Aristotle/De-anima/

Aristotle, more than any other ancient philosopher, set the terms of reference for the future discussion of the problem of consciousness. His idea of the mind is summarised in the illustration.

Aristotle was a physicalist, believing that things are embodied in the material universe:

"... That is precisely why the study of the soul must fall within the science of Nature, at least so far as in its affections it manifests this double character. Hence a physicist would define an affection of soul differently from a dialectician; the latter would define e.g. anger as the appetite for returning pain for pain, or something like that, while the former would define it as a boiling of the blood or warm substance surround the heart. The latter assigns the material conditions, the former the form or formulable essence; for what he states is the formulable essence of the fact, though for its actual existence there must be embodiment of it in a material such as is described by the other."(Book I)



The works of Aristotle provide our first clear account of the concept of signals and information. He was aware that an event can change the state of matter and this change of state can be transmitted to other locations where it can further change a state of matter:

"If what has colour is placed in immediate contact with the eye, it cannot be seen. Colour sets in movement not the sense organ but what is transparent, e.g. the air, and that, extending continuously from the object to the organ, sets the latter in movement. Democritus misrepresents the facts when he expresses the opinion that if the interspace were empty one could distinctly see an ant on the

vault of the sky; that is an impossibility. Seeing is due to an affection or change of what has the perceptive faculty, and it cannot be affected by the seen colour itself; it remains that it must be affected by what comes between. Hence it is indispensable that there be something in between-if there were nothing, so far from seeing with greater distinctness, we should see nothing at all." (Book II)

He was also clear about the relationship of information to 'state':

"By a 'sense' is meant what has the power of receiving into itself the sensible forms of things without the matter. This must be conceived of as taking place in the way in which a piece of wax takes on the impress of a signet-ring without the iron or gold; we say that what produces the impression is a signet of bronze or gold, but its particular metallic constitution makes no difference: in a similar way the sense is affected by what is coloured or flavoured or sounding, but it is indifferent what in each case the substance is; what alone matters is what quality it has, i.e. in what ratio its constituents are combined"(Book II)

Aristotle also mentioned the problem of the simultaneity of experience. The explanation predates Galilean and modern physics so lacks our modern language to explain how many things could be at a point and an instant:

"... just as what is called a 'point' is, as being at once one and two, properly said to be divisible, so here, that which discriminates is qua undivided one, and active in a single moment of time, while so far forth as it is divisible it twice over uses the same dot at one and the same time. So far forth then as it takes the limit as two' it discriminates two separate objects with what in a sense is divided: while so far as it takes it as one, it does so with what is one and occupies in its activity a single moment of time. (Book III)

He described the problem of recursion that would occur if the mind were due to the flow of material things in space:

"...mind is either without parts or is continuous in some other way than that which characterizes a spatial magnitude. How, indeed, if it were a spatial magnitude, could mind possibly think? Will it think with any one indifferently of its parts? In this case, the 'part' must be understood either in the sense of a spatial magnitude or in the sense of a point (if a point can be called a part of a spatial magnitude). If we accept the latter alternative, the points being infinite in number, obviously the mind can never exhaustively traverse them; if the former, the mind must think the same thing over and over again, indeed an infinite number of times (whereas it is manifestly possible to think a thing once only)."(Book I)

Aristotle explicitly mentions the regress:

"..we must fall into an infinite regress or we must assume a sense which is aware of itself." (Book III,425b)

However, this regress was not as problematic for Aristotle as it is for philosophers who are steeped in nineteenth century ideas. Aristotle was a physicalist who was not burdened with materialism and so was able to escape from the idea that the only possibility for the mind is a flow of material from place to place over a succession of disconnected instants. He was able to propose that subjects and objects are part of the same thing, he notes that thought is both temporally and spatially extended:

"But that which mind thinks and the time in which it thinks are in this case divisible only incidentally and not as such. For in them too there is something indivisible (though, it may be, not isolable) which gives unity to the time and the whole of length; and this is found equally in every continuum whether temporal or spatial."

This idea of time allowed him to identify thinking with the object of thought, there being no

need to cycle thoughts from instant to instant because mental time is extended:

"In every case the mind which is actively thinking is the objects which it thinks."

He considered imagination to be a disturbance of the sense organs:

"And because imaginations remain in the organs of sense and resemble sensations, animals in their actions are largely guided by them, some (i.e. the brutes) because of the non-existence in them of mind, others (i.e. men) because of the temporary eclipse in them of mind by feeling or disease or sleep.(Book III)"

And considered that all thought occurs as images:

"To the thinking soul images serve as if they were contents of perception (and when it asserts or denies them to be good or bad it avoids or pursues them). That is why the soul never thinks without an image."(Book III).

Aristotle also described the debate between the cognitive and behaviourist approaches with their overtones of the conflict between modern physicalism and pre twentieth century materialism:

"Some thinkers, accepting both premisses, viz. that the soul is both originative of movement and cognitive, have compounded it of both and declared the soul to be a self-moving number."(Book I)

The idea of a 'self-moving number' is not as absurd as it seems, like much of Ancient Greek philosophy.

Aristotle was also clear about there being two forms involved in perception. He proposed that the form and properties of the things that are directly in the mind are incontrovertible but that our inferences about the form and properties of the things in the world that give rise to the things in the mind can be false:

"Perception (1) of the special objects of sense is never in error or admits the least possible amount of falsehood. (2) That of the concomitance of the objects concomitant with the sensible qualities comes next: in this case certainly we may be deceived; for while the perception that there is white before us cannot be false, the perception that what is white is this or that may be false. (3) Third comes the perception of the universal attributes which accompany the concomitant objects to which the special sensibles attach (I mean e.g. of movement and magnitude); it is in respect of these that the greatest amount of sense-illusion is possible."(Book III)

Imagination, according to this model, lays out things in the senses.

Homer,(c.800-900 BC)The Iliad and Odyssey

Odyssey

Panpsychism and panexperientialism can be traced to, at least, Homer's Iliad. Just reading the book allows us to experience what a different focus of consciousness feels like. It is a way of being, Being an Homeric Greek, distinct from being a modern man. Both states of consciousness result in different ways of experiencing the world.

As we read the Iliad, we are drawn into the book through the images it creates in us and the feelings it evokes in us through the meter and the language. The reader becomes the book. 'The reader became the book, and the summer night was like the conscious being of the book' (Wallace Stevens). That experience of becoming the book, of loosing yourself in the book, is the experience of a different aspect of consciousness, being an Homeric Greek.

Homer frequently ascribes even our emotions to the world around us. The ancients do not just fear but fear grips them, for example: "So spake Athene, and pale fear gat hold of them all. The arms flew from their hands in their terror and fell all upon the ground, as the goddess uttered her voice" (Odyssey book XXIV).

The German classicist Bruno Snell, in 'The Discovery of the Mind' provides us with 'a convincing account of the enormous change in... human personality which took place during the centuries covered by Homer (to) Socrates.'(The London Times Literary Supplement). Snells book establishes two disinct aspects of consciousness. He says 'The experience of Homer differs from our own'(p.v). 'For Homer, psyche is the force which keeps the human being alive'(p.8). When the psyche leaves, the owner loses consciousness. The Homeric 'psyche' is where pan-pychism originates. It begins in a conception of consciousness as a force that is seperate from the body. Snell compares Homer to the tragedy of Orestes, which focuses on the individual. Homer concentrates on the action(process) and the situation in preference to the agent'(p.211) Orestes is in a different state of consciousness, 'a new state of consciousness'(p.211).

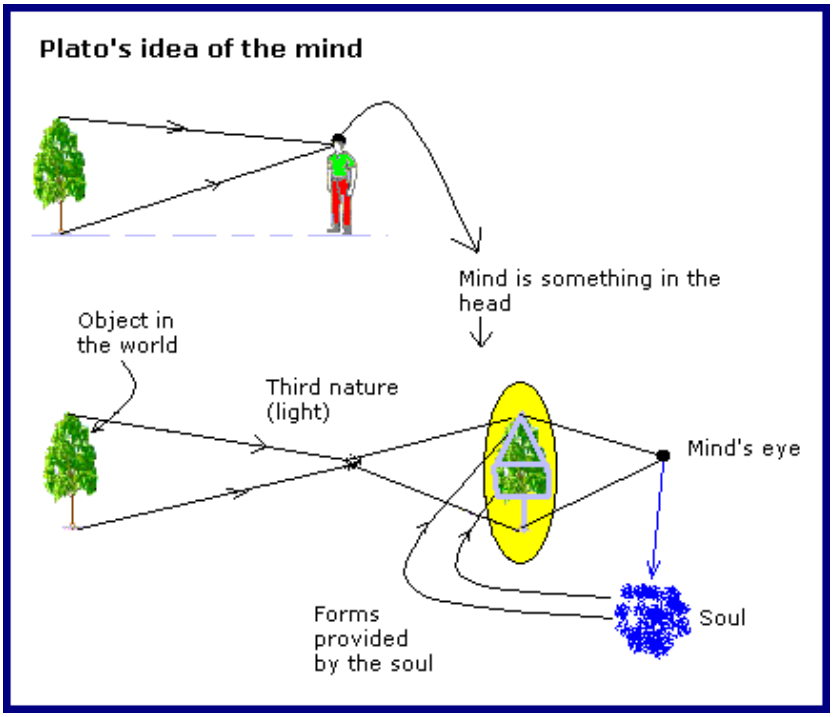
Plato (427-347BC)

The Republic http://www.constitution.org/pla/repub_00.htm Especially book VI http://www.constitution.org/pla/repub_06.htm

Plato's most interesting contributions to consciousness studies are in book VI of *The Republic*. His idea of the mind is illustrated below.

He believes that light activates pre-existing capabilities in the eyes:

"Sight being, as I conceive, in the eyes, and he who has eyes wanting to see; color being also present in them, still unless there be a third nature specially adapted to the purpose, the owner of the eyes will see nothing and the colors will be invisible."



However, it is in the metaphor of **the divided line** that Plato introduces a fascinating account of the relationships and properties of things. He points out that analysis deals in terms of the relationships of pure forms:

"And do you not know also that although they make use of the visible forms and reason about them, they are thinking not of these, but of the ideals which they resemble; not of the figures which they draw, but of the absolute square and the absolute diameter, and so on -- the forms which they draw or make, and which have shadows and reflections in water of their own, are converted by them into images, but they are really seeking to behold the things themselves, which can only be seen with the eye of the mind?"

Notice how he introduces the notion of a mind's eye observing mental content arranged as geometrical forms. He proposes that through this mode of ideas we gain understanding:

"And the habit which is concerned with geometry and the cognate sciences I suppose that you would term understanding, and not reason, as being intermediate between opinion and reason."

However, the understanding can also contemplate knowledge:

"..I understand you to say that knowledge and being, which the science of dialectic contemplates, are clearer than the notions of the arts, as they are termed, which proceed from hypotheses only: these are also contemplated by the understanding, and not by the senses: yet, because they start from hypotheses and do not ascend to a principle, those who contemplate them appear to you not to exercise the higher reason upon them, although when a first principle is added to them they are cognizable by the higher reason. "

Plato's work is not usually discussed in this way but is extended to universals such as the

idea of the colour red as a universal that can be applied to many specific instances of things.

Siddhartha Gautama c.500BC Buddhist Texts

Siddhartha Gautama was born about 563BC. He became known as 'Buddha' ('the awakened one') from the age of about thirty five. Buddha handed down a way of life that might lead, eventually, to an enlightened state called Nirvana. In the three centuries after his death Buddhism split into two factions, the Mahayana (greater raft or vehicle) and the Theravada (the way of the elders). The Mahayana use the slightly derogatory term Hinayana (lesser raft or vehicle) for Theravada Buddhism. Mahayana Buddhism gave rise to other sects such as Zen Buddhism in Japan and Vajrayana Buddhism in Tibet. Mahayana Buddhism is more like a religion, complete with god like entities whereas Theravada Buddhism is more like a philosophy.

Theravada Buddhist meditation is described in books called the Pali Canon which contains the 'Vinayas' that describe monastic life, the 'Suttas' which are the central teachings of Theravada Buddhism and the 'Abhidhamma' which is an analysis of the other two parts or 'pitakas'. Two meditational systems are described: the development of serenity (samathabhavana) and the development of insight (vipassanabhavana). The two systems are complementary, serenity meditation providing a steady foundation for the development of insight. As meditation proceeds the practitioner passes through a series of stages called 'jhanas'. There are four of these stages of meditation and then a final stage known as the stage of the 'immaterial jhanas'.

The Jhanas

The first jhana is a stage of preparation where the meditator rids themselves of the hindrances (sensual desire, ill will, sloth and torpor, restlessness and worry, and doubt). This is best achieved by seclusion. During the process of getting rid of the hindrances the meditator develops the five factors: applied thought, sustained thought, rapture, happiness and one-pointedness of mind. This is done by concentrating on a practice object until it can be easily visualised. Eventually the mediator experiences a luminous replica of the object called the counterpart sign (patibhaganimitta).

Applied thought involves examining, visualising and thinking about the object. Sustained thought involves always returning to the object, not drifting away from it. Rapture involves a oneness with the object and is an ecstacy that helps absorption with and in the object. Happiness is the feeling of happiness that everyone has when something good happens (unlike rapture, which is a oneness with the object of contemplation). One-pointedness of mind is the ability to focus on a single thing without being distracted.

The second jhana involves attaining the first without effort, there is no need for applied or sustained thought, only rapture, happiness and one-pointedness of mind remain. The second jhana is achieved by contemplating the first jhana. The second jhana is a stage of effortless concentration.

The third jhana involves mindfulness and discernment. The mindfulness allows an object of meditation to be held effortlessly in the mind. The discernment consists of discerning the nature of the object without delusion and hence avoiding rapture.

In the fourth jhana mindfulness is maintained but the delusion of happiness is contemplated. Eventually mindfulness remains without pleasure or pain. In the fourth jhana the meditator achieves "purity of mindfulness due to equanimity" (upekkhasatiparisuddhi).

The Immaterial Jhanas

The first four jhanas will be familiar from earlier, Hindu meditational techniques. Once the fourth jhana has been achieved the meditator can embark on the immaterial jhanas. There are four immaterial jhanas: the base of boundless space, the base of boundless consciousness, the base of nothingness, and the base of neither-perception-nor-non-perception.

The base of boundless space is achieved by meditating on the absence of the meditation object. It is realised that the space occupied by the object is boundless and that the mind too is boundless space. The base of boundless consciousness involves a realisation that the boundless space is boundless consciousness. The base of nothingness is a realisation that the present does not exist, the meditator should "give attention to the present non-existence, voidness, secluded aspect of that same past consciousness belonging to the base consisting of boundless space" (Gunaratana 1988). The base of neither-perception-nor-non-perception is a realisation that nothing is perceived in the void.

In Theravada Buddhism the attainment of the fourth jhana and its immaterial jhanas represents a mastery of serenity meditation. This is a foundation for insight meditation.

Buddhism is very practical and eschews delusions. It is realised that serenity meditation is a state of mind, a steady foundation that might, nowadays be called a physiological state. It is through insight meditation where the practitioner becomes a philosopher that enlightenment is obtained.

Further reading:

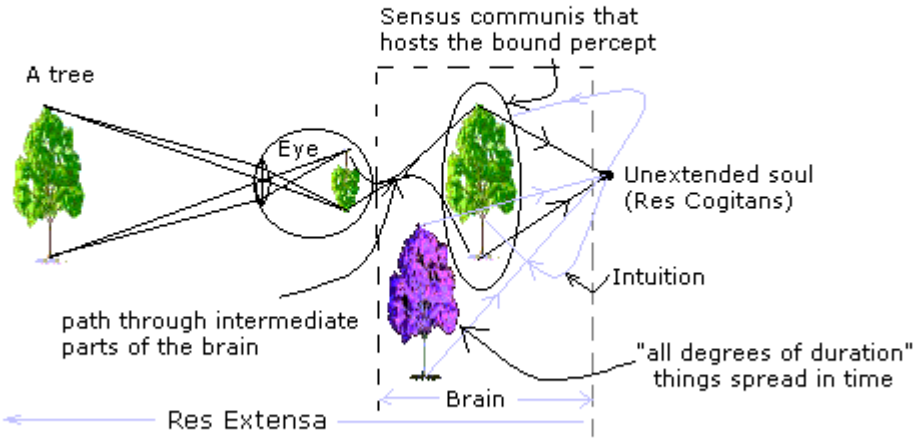
The Buddhist Publication Society. Especially: The Jhanas In Theravada Buddhist Meditation by Henepola Gunaratana. The Wheel Publication No. 351/353 ISBN 955-24-0035-X. 1988 Buddhist Publication Society. http://www.accesstoinsight.org/lib/bps/index.html

Seventeenth and Eighteenth Century Philosophy

Rene Descartes (1596-1650)

Descartes was also known as Cartesius. He had an empirical approach to consciousness and the mind, describing in his *Meditations on First Philosophy* (1641) what it is like to be human. His idea of perception is summarised in the diagram below.

Descartes idea of perception



Dubitability

Descartes is probably most famous for his statement:

"But immediately upon this I observed that, whilst I thus wished to think that all was false, it was absolutely necessary that I, who thus thought, should be somewhat; and as I observed that this truth, I think, therefore I am (COGITO ERGO SUM), was so certain and of such evidence that no ground of doubt, however extravagant, could be alleged by the sceptics capable of shaking it, I concluded that I might, without scruple, accept it as the first principle of the philosophy of which I was in search."

Descartes is clear that what he means by *thought* is all the things that occur in experience, whether dreams, sensations, symbols etc.:

"5. Of my thoughts some are, as it were, images of things, and to these alone properly belongs the name IDEA; as when I think [represent to my mind] a man, a chimera, the sky, an angel or God. Others, again, have certain other forms; as when I will, fear, affirm, or deny, I always, indeed, apprehend something as the object of my thought, but I also embrace in thought something more than the representation of the object; and of this class of thoughts some are called volitions or affections, and others judgments." (Meditation III).

He repeats this general description of thought in many places in the Meditations and elsewhere. What Descartes is saying is that his meditator has thoughts; that there are

thoughts and this cannot be doubted when and where they occur (Russell (1945) makes this clear).

Needless to say the basic *cogito* put forward by Descartes has provoked endless debate, much of it based on the false premise that Descartes was presenting an inference or argument rather than just saying that thought certainly exists. However, the extent to which the philosopher can go beyond this certainty to concepts such as God, science or the soul is highly problematical.

The description of thoughts and mind

Descartes uses the words "ideas" and "imagination" in a rather unusual fashion. The word "idea" he defines as follows:

"5. Of my thoughts some are, as it were, images of things, and to these alone properly belongs the name IDEA; as when I think [represent to my mind] a man, a chimera, the sky, an angel or God. Others, again, have certain other forms; as when I will, fear, affirm, or deny, I always, indeed, apprehend something as the object of my thought, but I also embrace in thought something more than the representation of the object; and of this class of thoughts some are called volitions or affections, and others judgments." (Meditation III).

As will be seen later, Descartes regards his mind as an unextended thing (a point) so "images of things" or "IDEAS" require some way of being extended. In the *Treatise on Man* (see below) he is explicit that ideas are extended things in the brain, on the surface of the "common sense". In *Rules for the Direction of the Mind* he notes that we "receive ideas from the common sensibility", an extended part of the brain. This usage of the term "ideas" is very strange to the modern reader and the source of many mistaken interpretations. It should be noted that occasionally Descartes uses the term 'idea' according to its usual meaning where it is almost interchangeable with 'thought' in general but usually he means a representation laid out in the brain.

Descartes considers the imagination to be the way that the mind "turns towards the body" (by which Descartes means the part of the brain in the body called the senses communis):

"3. I remark, besides, that this power of imagination which I possess, in as far as it differs from the power of conceiving, is in no way necessary to my [nature or] essence, that is, to the essence of my mind; for although I did not possess it, I should still remain the same that I now am, from which it seems we may conclude that it depends on something different from the mind. And I easily understand that, if some body exists, with which my mind is so conjoined and united as to be able, as it were, to consider it when it chooses, it may thus imagine corporeal objects; so that this mode of thinking differs from pure intellection only in this respect, that the mind in conceiving turns in some way upon itself, and considers some one of the ideas it possesses within itself; but in imagining it turns toward the body, and contemplates in it some object conformed to the idea which it either of itself conceived or apprehended by sense." Meditations VI

So ideas, where they become imagined images of things were thought by Descartes to involve a phase of creating a form in the brain.

Descartes gives a clear description of his experience as a container that allows length, breadth, depth, continuity and time with contents arranged within it:

"2. But before considering whether such objects as I conceive exist without me, I must examine their ideas in so far as these are to be found in my consciousness, and discover which of them are distinct and which confused.

3. In the first place, I distinctly imagine that quantity which the philosophers commonly call continuous, or the extension in length, breadth, and depth that is in this quantity, or rather in the object to which it is attributed. Further, I can enumerate in it many diverse parts, and attribute to each of these all sorts of sizes, figures, situations, and local motions; and, in fine, I can assign to each of these motions all degrees of duration."(Meditation V).

He points out that sensation occurs by way of the brain, conceptualising the brain as the place in the body where the extended experiences are found : Meditations VI:

"20. I remark, in the next place, that the mind does not immediately receive the impression from all the parts of the body, but only from the brain, or perhaps even from one small part of it, viz., that in which the common sense (senses communis) is said to be, which as often as it is affected in the same way gives rise to the same perception in the mind, although meanwhile the other parts of the body may be diversely disposed, as is proved by innumerable experiments, which it is unnecessary here to enumerate."

He finds that both imaginings and perceptions are extended things and hence in the (brain part) of the body. The area of extended things is called the *res extensa*, it includes the brain, body and world beyond. He also considers the origin of intuitions, suggesting that they can enter the mind without being consciously created: Meditations VI, 10 :

"10. Moreover, I find in myself diverse faculties of thinking that have each their special mode: for example, I find I possess the faculties of imagining and perceiving, without which I can indeed clearly and distinctly conceive myself as entire, but I cannot reciprocally conceive them without conceiving myself, that is to say, without an intelligent substance in which they reside, for [in the notion we have of them, or to use the terms of the schools] in their formal concept, they comprise some sort of intellection; whence I perceive that they are distinct from myself as modes are from things. I remark likewise certain other faculties, as the power of changing place, of assuming diverse figures, and the like, that cannot be conceived and cannot therefore exist, any more than the preceding, apart from a substance in which they inhere. It is very evident, however, that these faculties, if they really exist, must belong to some corporeal or extended substance, since in their clear and distinct concept there is contained some sort of extension, but no intellection at all. Further, I cannot doubt but that there is in me a certain passive faculty of perception, that is, of receiving and taking knowledge of the ideas of sensible things; but this would be useless to me, if there did not also exist in me, or in some other thing, another active faculty capable of forming and producing those ideas. But this active faculty cannot be in me [in as far as I am but a thinking thing], seeing that it does not presuppose thought, and also that those ideas are frequently produced in my mind without my contributing to it in any way, and even frequently contrary to my will. This faculty must therefore exist in some substance different from me, in which all the objective reality of the ideas that are produced by this faculty is contained formally or eminently, as I before remarked; and this substance is either a body, that is to say, a corporeal nature in which is contained formally [and in effect] all that is objectively [and by representation] in those ideas; or it is God himself, or some other creature, of a rank superior to body, in which the same is contained eminently. But as God is no deceiver, it is manifest that he does not of himself and immediately communicate those ideas to me, nor even by the intervention of any creature in which their objective reality is not formally, but only eminently, contained. For as he has given me no faculty whereby I can discover this to be the case, but, on the contrary, a very strong inclination to believe that those ideas arise from corporeal objects, I do not see how he could be vindicated from the charge of deceit, if in truth they proceeded from any other source, or were produced by other causes than corporeal things: and accordingly it must be concluded, that corporeal objects exist. Nevertheless, they are not perhaps exactly such as we perceive by the senses, for their comprehension by the senses is, in many instances, very obscure and confused; but it is at least necessary to admit that all which I clearly and distinctly conceive as in them, that is, generally speaking all that is comprehended in the object of speculative geometry, really exists external to me. "

He considers that the mind itself is the thing that generates thoughts and is not extended (occupies no space). This 'mind' is known as the *res cogitans*. The mind works on the

imaginings and perceptions that exist in that part of the body called the brain. This is Descartes' dualism: it is the proposition that there is an unextended place called the mind that acts upon the extended things in the brain. Meditations VI, 9:

"... And although I may, or rather, as I will shortly say, although I certainly do possess a body with which I am very closely conjoined; nevertheless, because, on the one hand, I have a clear and distinct idea of myself, in as far as I am only a thinking and unextended thing, and as, on the other hand, I possess a distinct idea of body, in as far as it is only an extended and unthinking thing, it is certain that I, [that is, my mind, by which I am what I am], is entirely and truly distinct from my body, and may exist without it."

Notice that the intellection associated with ideas is part of an "active faculty capable of forming and producing those ideas" that has a "corporeal nature" (it is in the brain). This suggests that the "thinking" in the passage above applies only to those thoughts that are unextended, however, it is difficult to find a definition of these particular thoughts.

"Rules for the Direction of the Mind" demonstrates Descartes' dualism. He describes the brain as the part of the body that contains images or phantasies of the world but believes that there is a further, spiritual mind that processes the images in the brain:

"My fourth supposition is that the power of movement, in fact the nerves, originate in the brain, where the phantasy is seated; and that the phantasy moves them in various ways, as the external sense <organ> moves the <organ of> common sensibility, or as the whole pen is moved by its tip. This illustration also shows how it is that the phantasy can cause various movements in the nerves, although it has not images of these formed in itself, but certain other images, of which these movements are possible effects. For the pen as a whole does not move in the same way as its tip; indeed, the greater part of the pen seems to go along with an altogether different, contrary motion. This enables us to understand how the movements of all other animals are accomplished, although we suppose them to have no consciousness (rerum cognitio) but only a bodily <organ of> phantasy; and furthermore, how it is that in ourselves those operations are performed which occur without any aid of reason.

My fifth and last supposition is that the power of cognition properly so called is purely spiritual, and is just as distinct from the body as a whole as blood is from bone or a hand from an eye; and that it is a single power. Sometimes it receives images from the common sensibility at the same time as the phantasy does; sometimes it applies itself to the images preserved in memory; sometimes it forms new images, and these so occupy the imagination that often it is not able at the same time to receive ideas from the common sensibility, or to pass them on to the locomotive power in the way that the body left to itself -would. "

Descartes sums up his concept of a point soul seeing forms in the world via forms in the sensus communis in *Passions of the Soul*, 35:

"By this means the two images which are in the brain form but one upon the gland, which, acting immediately upon the soul, causes it to see the form in the mind".

Anatomical and physiological ideas

In his *Treatise on Man* Descartes summarises his ideas on how we perceive and react to things as well as how consciousness is achieved anatomically and physiologically. The 'Treatise' was written at a time when even galvanic electricity was unknown. The excerpt given below covers Descartes' analysis of perception and stimulus-response processing.

Fig. 1

"Thus for example [in Fig 1], if fire A is close to foot B, the tiny parts of this fire (which, as you know, move about very rapidly) have the power also to move the area of skin which they touch. In this way they pull the tiny fibre *cc* which you see attached to it, and simultaneously open the entrance to the pore *de*, located opposite the point where this fiber terminates - just as when you pull one end of a string, you cause a bell hanging at the other end to ring at the same time.

When the entrance to the pore or small tube *de* is opened in this way, the animal spirits from cavity F enter and are carried through it - some to muscles which serve to pull the foot away from the fire, some to muscles which turn the eyes and head to look at it, and some to muscles which make the hands move and the whole body turn in order to protect it.

Now I maintain that when God unites a rational soul to this machine (in a way that I intend to explain later) he will place its principle seat in the brain, and will make its nature such that the soul will have different sensations corresponding to the different ways in which the entrances to the pores in the internal surface of the brain are opened by means of nerves.

In order to see clearly how ideas are formed of the objects which strike the senses, observe in this diagram [fig 2] the tiny fibres 12, 34, 56, and the like, which make up the optic nerve and stretch from the back of the eye at 1, 3, 5 to the internal surface of the brain at 2, 4, 6. Now assume that these fibres are so arranged that if the rays coming, for example, from point A of the object happen to press upon the back of the eye at point 1, they pull the whole of fibre 12 and enlarge the opening of the tiny tube marked 2. In the same way, the rays which come from point B enlarge the opening of the tiny tube 4, and likewise for the others.

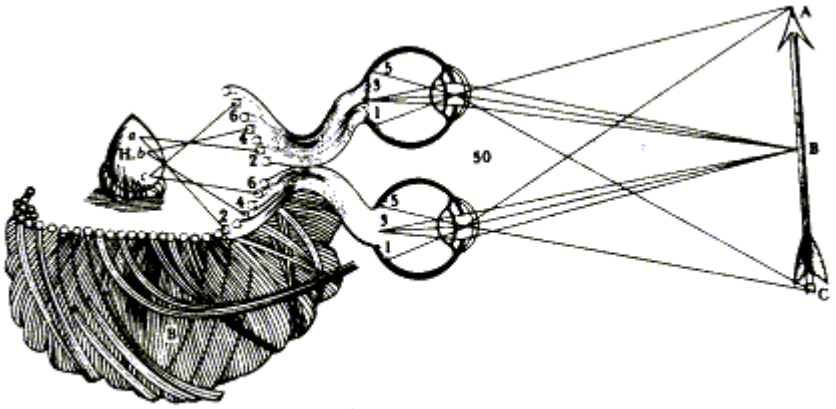


Fig. 2

We have already described how, depending on the different ways in which the points 1, 3, 5 are pressed by these rays, a figure is traced on the back of the eye corresponding to that of the object ABC. Similarly it is obvious that, depending on the different ways in which the tiny tubes 2, 4, 6 are opened by the fibres 12, 34, 56 etc., a corresponding figure must also be traced on the internal surface of the brain.

.....

And note that by 'figures' I mean not only things which somehow represent the position of the edges and surfaces of objects, but also anything which, as I said above, can give the soul occasion to perceive movement, size, distance, colours, sounds, smells and other such qualities. And I also include anything that can make the sould feel pleasure, pain, hunger, thirts, joy, sadness and other such passions.

...

Now among these figures, it is not those imprinted on the external sense organs, or on the internal surface of the brain, which should be taken to be ideas - but only those which are traced in the spirits on the surface of gland H (*where the seat of the imagination and the 'common sense' is located*). That is to say, it is only the latter figures which should be taken to be the forms or images which the rational soul united to this machine will consider directly when it imagines some object or perceives it by the senses.

And note that I say 'imagines or perceives by the senses'. For I wish to apply the term 'idea' generally to all impressions which the spirits can receive as they leave gland H. These are to be attributed to the 'common' sense when they depend on the presence of objects; but they may also proceed from many other causes (as I shall explain later), and they should then be attributed to the imagination. "

The *common sense* is referred to by philosophers as the *senses communis*. Descartes considered this to be the place where all the sensations were bound together and proposed the pineal gland for this role. This was in the days before the concept of 'dominance' of parts of the brain had been developed so Descartes reasoned that only a single organ could host a bound representation.

Notice how Descartes is explicit about ideas being **traced in the spirits on the surface of the gland**. Notice also how the rational soul will consider forms on the common sense **directly**.

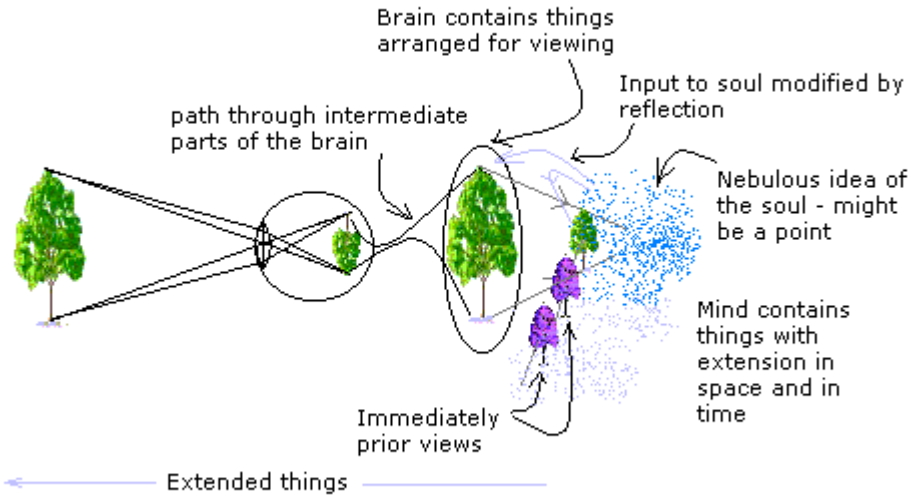
Descartes believed that animals are not conscious because, although he thought they possessed the stimulus-response loop in the same way as humans he believed that they do

not possess a soul.

John Locke (1632-1704)

Locke's most important philosophical work on the human mind was "An Essay Concerning Human Understanding" written in 1689. His idea of perception is summarised in the diagram below:

Locke's idea of perception



Locke is an Indirect Realist, admitting of external objects but describing these as represented within the mind. The objects themselves are thought to have a form and properties that are the *archetype* of the object and these give rise in the brain and mind to derived copies called *ektypa*.

Like Descartes, he believes that people have souls that produce thoughts. Locke considers that sensations make their way from the senses to the brain where they are laid out for understanding as a 'view':

"And if these organs, or the nerves which are the conduits to convey them from without to their audience in the brain,- the mind's presence-room (as I may so call it)- are any of them so disordered as not to perform their functions, they have no postern to be admitted by; no other way to bring themselves into view, and be perceived by the understanding." (Chapter III, 1).

He considers that what is sensed becomes a mental thing: Chapter IX: Of Perception paragraph 1:

"This is certain, that whatever alterations are made in the body, if they reach not the mind; whatever impressions are made on the outward parts, if they are not taken notice of within, there is no perception. Fire may burn our bodies with no other effect than it does a billet, unless the motion be continued to the brain, and there the sense of heat, or idea of pain, be produced in the mind;

wherein consists actual perception. "

Locke calls the contents of consciousness "ideas" (cf: Descartes, Malebranche) and regards sensation, imagination etc. as being similar or even alike. Chapter I: Of Ideas in general, and their Original:

"1. Idea is the object of thinking. Every man being conscious to himself that he thinks; and that which his mind is applied about whilst thinking being the ideas that are there, it is past doubt that men have in their minds several ideas,- such as are those expressed by the words whiteness, hardness, sweetness, thinking, motion, man, elephant, army, drunkenness, and others: it is in the first place then to be inquired, How he comes by them?

I know it is a received doctrine, that men have native ideas, and original characters, stamped upon their minds in their very first being. This opinion I have at large examined already; and, I suppose what I have said in the foregoing Book will be much more easily admitted, when I have shown whence the understanding may get all the ideas it has; and by what ways and degrees they may come into the mind;- for which I shall appeal to every one's own observation and experience.

2. All ideas come from sensation or reflection. Let us then suppose the mind to be, as we say, white paper, void of all characters, without any ideas:- How comes it to be furnished? Whence comes it by that vast store which the busy and boundless fancy of man has painted on it with an almost endless variety? Whence has it all the materials of reason and knowledge? To this I answer, in one word, from EXPERIENCE. In that all our knowledge is founded; and from that it ultimately derives itself. Our observation employed either, about external sensible objects, or about the internal operations of our minds perceived and reflected on by ourselves, is that which supplies our understandings with all the materials of thinking. These two are the fountains of knowledge, from whence all the ideas we have, or can naturally have, do spring.

3. The objects of sensation one source of ideas. First, our Senses, conversant about particular sensible objects, do convey into the mind several distinct perceptions of things, according to those various ways wherein those objects do affect them. And thus we come by those ideas we have of yellow, white, heat, cold, soft, hard, bitter, sweet, and all those which we call sensible qualities; which when I say the senses convey into the mind, I mean, they from external objects convey into the mind what produces there those perceptions. This great source of most of the ideas we have, depending wholly upon our senses, and derived by them to the understanding, I call SENSATION.

4. The operations of our minds, the other source of them. Secondly, the other fountain from which experience furnisheth the understanding with ideas is,- the perception of the operations of our own mind within us, as it is employed about the ideas it has got;- which operations, when the soul comes to reflect on and consider, do furnish the understanding with another set of ideas, which could not be had from things without. And such are perception, thinking, doubting, believing, reasoning, knowing, willing, and all the different actings of our own minds;- which we being conscious of, and observing in ourselves, do from these receive into our understandings as distinct ideas as we do from bodies affecting our senses. This source of ideas every man has wholly in himself; and though it be not sense, as having nothing to do with external objects, yet it is very like it, and might properly enough be called internal sense. But as I call the other SENSATION, so I Call this REFLECTION, the ideas it affords being such only as the mind gets by reflecting on its own operations within itself. By reflection then, in the following part of this discourse, I would be understood to mean, that notice which the mind takes of its own operations, and the manner of them, by reason whereof there come to be ideas of these operations in the understanding. These two, I say, viz. external material things, as the objects of SENSATION, and the operations of our own minds within, as the objects of REFLECTION, are to me the only originals from whence all our ideas take their beginnings. The term operations here I use in a large sense, as comprehending not barely the actions of the mind about its ideas, but some sort of passions arising sometimes from them, such as is the satisfaction or uneasiness arising from any thought.

5. All our ideas are of the one or the other of these. The understanding seems to me not to have the least glimmering of any ideas which it doth not receive from one of these two. External objects

furnish the mind with the ideas of sensible qualities, which are all those different perceptions they produce in us; and the mind furnishes the understanding with ideas of its own operations. "

He calls ideas that come directly from the senses *primary qualities* and those that come from reflection upon these he calls *secondary qualities*:

"9. Primary qualities of bodies. Qualities thus considered in bodies are, First, such as are utterly inseparable from the body, in what state soever it be; and such as in all the alterations and changes it suffers, all the force can be used upon it, it constantly keeps; and such as sense constantly finds in every particle of matter which has bulk enough to be perceived; and the mind finds inseparable from every particle of matter, though less than to make itself singly be perceived by our senses: ..... These I call original or primary qualities of body, which I think we may observe to produce simple ideas in us, viz. solidity, extension, figure, motion or rest, and number. 10. Secondary qualities of bodies. Secondly, such qualities which in truth are nothing in the objects themselves but power to produce various sensations in us by their primary qualities....." (Chapter VIII).

He gives examples of secondary qualities:

"13. How secondary qualities produce their ideas. After the same manner, that the ideas of these original qualities are produced in us, we may conceive that the ideas of secondary qualities are also produced, viz. by the operation of insensible particles on our senses. .....v.g. that a violet, by the impulse of such insensible particles of matter, of peculiar figures and bulks, and in different degrees and modifications of their motions, causes the ideas of the blue colour, and sweet scent of that flower to be produced in our minds. It being no more impossible to conceive that God should annex such ideas to such motions, with which they have no similitude, than that he should annex the idea of pain to the motion of a piece of steel dividing our flesh, with which that idea hath no resemblance." (Chapter VIII).

He argues against all conscious experience being in mental space (does not consider that taste might be on the tongue or a smell come from a cheese): Chapter XIII: Complex Ideas of Simple Modes:- and First, of the Simple Modes of the Idea of Space - paragraph 25:

"I shall not now argue with those men, who take the measure and possibility of all being only from their narrow and gross imaginations: but having here to do only with those who conclude the essence of body to be extension, because they say they cannot imagine any sensible quality of any body without extension,- I shall desire them to consider, that, had they reflected on their ideas of tastes and smells as much as on those of sight and touch; nay, had they examined their ideas of hunger and thirst, and several other pains, they would have found that they included in them no idea of extension at all, which is but an affection of body, as well as the rest, discoverable by our senses, which are scarce acute enough to look into the pure essences of things."

Locke understood the "specious" or extended present but conflates this with longer periods of time: Chapter XIV. Idea of Duration and its Simple Modes - paragraph 1:

"Duration is fleeting extension. There is another sort of distance, or length, the idea whereof we get not from the permanent parts of space, but from the fleeting and perpetually perishing parts of succession. This we call duration; the simple modes whereof are any different lengths of it whereof we have distinct ideas, as hours, days, years, &c., time and eternity."

Locke is uncertain about whether extended ideas are viewed from an unextended soul.

"He that considers how hardly sensation is, in our thoughts, reconcilable to extended matter; or existence to anything that has no extension at all, will confess that he is very far from certainly knowing what his soul is. It is a point which seems to me to be put out of the reach of our knowledge: and he who will give himself leave to consider freely, and look into the dark and intricate part of each hypothesis, will scarce find his reason able to determine him fixedly for or against the soul's materiality. Since, on which side soever he views it, either as an unextended substance, or as a thinking extended matter, the difficulty to conceive either will, whilst either alone is in his thoughts,

still drive him to the contrary side."(Chapter III, 6).

David Hume (1711-1776)

Hume (1739-40). A Treatise of Human Nature: Being An Attempt to Introduce the Experimental Method of Reasoning Into Moral Subjects. http://www.class.uidaho.edu/mickelsen/ToC/hume%20treatise%20ToC.htm

Hume represents a type of pure empiricism where certainty is only assigned to present experience. As we can only directly know the mind he works within this constraint. He admits that there can be consistent bodies of knowledge within experience and would probably regard himself as an Indirect Realist but with the caveat that the things that are inferred to be outside the mind, in the physical world, could be no more than inferences within the mind.

Hume has a clear concept of mental space and time that is informed by the senses:

"The idea of space is convey'd to the mind by two senses, the sight and touch; nor does anything ever appear extended, that is not either visible or tangible. That compound impression, which represents extension, consists of several lesser impressions, that are indivisible to the eye or feeling, and may be call'd impressions of atoms or corpuscles endow'd with colour and solidity. But this is not all. 'Tis not only requisite, that these atoms shou'd be colour'd or tangible, in order to discover themselves to our senses; 'tis also necessary we shou'd preserve the idea of their colour or tangibility in order to comprehend them by our imagination. There is nothing but the idea of their colour or tangibility, which can render them conceivable by the mind. Upon the removal of the ideas of these sensible qualities, they are utterly annihilated to the thought or imagination.'

Now such as the parts are, such is the whole. If a point be not consider'd as colour'd or tangible, it can convey to us no idea; and consequently the idea of extension, which is compos'd of the ideas of these points, can never possibly exist. But if the idea of extension really can exist, as we are conscious it does, its parts must also exist; and in order to that, must be consider'd as colour'd or tangible. We have therefore no idea of space or extension, but when we regard it as an object either of our sight or feeling.

The same reasoning will prove, that the indivisible moments of time must be fill'd with some real object or existence, whose succession forms the duration, and makes it be conceivable by the mind."

In common with Locke and Eastern Philosophy, Hume considers reflection and sensation to be similar, perhaps identical:

"Thus it appears, that the belief or assent, which always attends the memory and senses, is nothing but the vivacity of those perceptions they present; and that this alone distinguishes them from the imagination. To believe is in this case to feel an immediate impression of the senses, or a repetition of that impression in the memory. 'Tis merely the force and liveliness of the perception, which constitutes the first act of the judgment, and lays the foundation of that reasoning, which we build upon it, when we trace the relation of cause and effect."

Hume considers that the origin of sensation can never be known, believing that the canvass of the mind contains our view of the world whatever the ultimate source of the images within the view and that we can construct consistent bodies of knowledge within these constraints:

"As to those impressions, which arise from the senses, their ultimate cause is, in my opinion, perfectly inexplicable by human reason, and 'twill always be impossible to decide with certainty, whether they arise immediately from the object, or are produc'd by the creative power of the mind, or

are deriv'd from the author of our being. Nor is such a question any way material to our present purpose. We may draw inferences from the coherence of our perceptions, whether they be true or false; whether they represent nature justly, or be mere illusions of the senses."

It may be possible to trace the origins of Jackson's Knowledge Argument in Hume's work:

" Suppose therefore a person to have enjoyed his sight for thirty years, and to have become perfectly well acquainted with colours of all kinds, excepting one particular shade of blue, for instance, which it never has been his fortune to meet with. Let all the different shades of that colour, except that single one, be plac'd before him, descending gradually from the deepest to the lightest; 'tis plain, that he will perceive a blank, where that shade is wanting, said will be sensible, that there is a greater distance in that place betwixt the contiguous colours, than in any other. Now I ask, whether 'tis possible for him, from his own imagination, to supply this deficiency, and raise up to himself the idea of that particular shade, tho' it had never been conveyed to him by his senses? I believe i here are few but will be of opinion that he can; and this may serve as a proof, that the simple ideas are not always derived from the correspondent impressions; tho' the instance is so particular and singular, that 'tis scarce worth our observing, and does not merit that for it alone we should alter our general maxim."

David Hume (1748) An Enquiry Concerning Human Understanding
http://www.infidels.org/library/historical/david_hume/human_understanding.html

Hume's view of Locke and Malebranche:

"The fame of Cicero flourishes at present; but that of Aristotle is utterly decayed. La Bruyere passes the seas, and still maintains his reputation: But the glory of Malebranche is confined to his own nation, and to his own age. And Addison, perhaps, will be read with pleasure, when Locke shall be entirely forgotten."

He is clear about relational knowledge in space and time:

"13. .. But though our thought seems to possess this unbounded liberty, we shall find, upon a nearer examination, that it is really confined within very narrow limits, and that all this creative power of the mind amounts to no more than the faculty of compounding, transposing, augmenting, or diminishing the materials afforded us by the senses and experience. When we think of a golden mountain, we only join two consistent ideas, gold, and mountain, with which we were formerly acquainted."

...

19. Though it be too obvious to escape observation, that different ideas are connected together; I do not find that any philosopher has attempted to enumerate or class all the principles of association; a subject, however, that seems worthy of curiosity. To me, there appear to be only three principles of connexion among ideas, namely, Resemblance, Contiguity in time or place, and Cause or Effect."

He is also clear that, although we experience the output of processes, we do not experience the processes themselves:

"29. It must certainly be allowed, that nature has kept us at a great distance from all her secrets, and has afforded us only the knowledge of a few superficial qualities of objects; while she conceals from us those powers and principles on which the influence of those objects entirely depends. Our senses inform us of the colour, weight, and consistence of bread; but neither sense nor reason can ever inform us of those qualities which fit it for the nourishment and support of a human body. Sight or feeling conveys an idea of the actual motion of bodies; but as to that wonderful force or power, which would carry on a moving body for ever in a continued change of place, and which bodies never lose but by communicating it to others; of this we cannot form the most distant conception. ..

58. ... All events seem entirely loose and separate. One event follows another; but we never can observe any tie between them. They seem conjoined, but never connected. And as we can have no idea of any thing which never appeared to our outward sense or inward sentiment, the necessary conclusion seems to be that we have no idea of connexion or power at all, and that these words are absolutely without any meaning, when employed either in philosophical reasonings or common life. "

Our idea of process is not a direct experience but seems to originate from remembering the repetition of events:

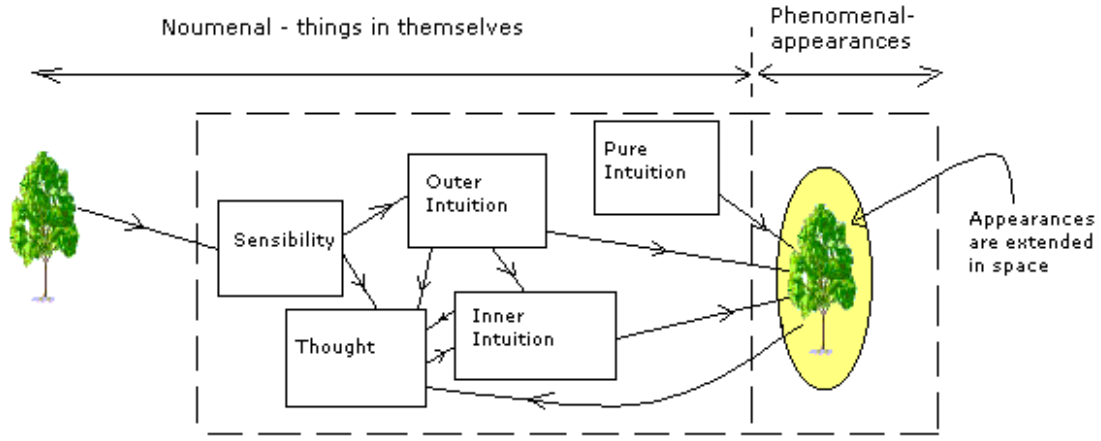
"59 ..It appears, then, that this idea of a necessary connexion among events arises from a number of similar instances which occur of the constant conjunction of these events; nor can that idea ever be suggested by any one of these instances, surveyed in all possible lights and positions. But there is nothing in a number of instances, different from every single instance, which is supposed to be exactly similar; except only, that after a repetition of similar instances, the mind is carried by habit, upon the appearance of one event, to expect its usual attendant, and to believe that it will exist."

Immanuel Kant (1724-1804)

Kant's greatest work on the subject of consciousness and the mind is Critique of Pure Reason (1781). Kant describes his objective in this work as discovering the axioms ("a priori concepts") and then the processes of 'understanding'.

P12 "This enquiry, which is somewhat deeply grounded, has two sides. The one refers to the objects of pure understanding, and is intended to expound and render intelligible the objective validity of its a priori concepts. It is therefore essential to my purposes. The other seeks to investigate the pure understanding itself, its possibility and the cognitive faculties upon which it rests; and so deals with it in its subjective aspect. Although this latter exposition is of great importance for my chief purpose, it does not form an essential part of it. For the chief question is always simply this: - what and how much can the understanding and reason know apart from all experience?"

Kant's idea of perception and mind



Kant's idea of perception and mind is summarised in the illustration below:

'Experience' is simply accepted. Kant believes that the physical world exists but is not

known directly:

P 24 "For we are brought to the conclusion that we can never transcend the limits of possible experience, though that is precisely what this science is concerned, above all else, to achieve. This situation yields, however, just the very experiment by which, indirectly, we are enabled to prove the truth of this first estimate of our a priori knowledge of reason, namely, that such knowledge has to do only with appearances, and must leave the thing in itself as indeed real per se, but as not known by us. "

Kant is clear about the form and content of conscious experience. He notes that we can only experience things that have appearance and 'form' - content and geometrical arrangement.

P65-66 "IN whatever manner and by whatever means a mode of knowledge may relate to objects, intuition is that through which it is in immediate relation to them, and to which all thought as a means is directed. But intuition takes place only in so far as the object is given to us. This again is only possible, to man at least, in so far as the mind is affected in a certain way. The capacity (receptivity) for receiving representations through the mode in which we are affected by objects, is entitled sensibility. Objects are given to us by means of sensibility, and it alone yields us intuitions; they are thought through the understanding, and from the understanding arise concepts. But all thought must, directly or indirectly, by way of certain characters relate ultimately to intuitions, and therefore, with us, to sensibility, because in no other way can an object be given to us. The effect of an object upon the faculty of representation, so far as we are affected by it, is sensation. That intuition which is in relation to the object through sensation, is entitled empirical. The undetermined object of an empirical intuition is entitled appearance. That in the appearance which corresponds to sensation I term its matter; but that which so determines the manifold of appearance that it allows of being ordered in certain relations, I term the form of appearance. That in which alone the sensations can be posited and ordered in a certain form, cannot itself be sensation; and therefore, while the matter of all appearance is given to us a posteriori only, its form must lie ready for the sensations a priori in the mind, and so must allow of being considered apart from all sensation. "

Furthermore he realises that experience exists without much content. That consciousness depends on form:

P66 "The pure form of sensible intuitions in general, in which all the manifold of intuition is intuited in certain relations, must be found in the mind a priori. This pure form of sensibility may also itself be called pure intuition. Thus, if I take away from the representation of a body that which the understanding thinks in regard to it, substance, force, divisibility, etc. , and likewise what belongs to sensation, impenetrability, hardness, colour, etc. , something still remains over from this empirical intuition, namely, extension and figure. These belong to pure intuition, which, even without any actual object of the senses or of sensation, exists in the mind a priori as a mere form of sensibility. The science of all principles of a priori sensibility I call transcendental aesthetic."

Kant proposes that space exists in our experience and that experience could not exist without it (apodeictic means 'incontrovertible):

P 68 "1. Space is not an empirical concept which has been derived from outer experiences. For in order that certain sensations be referred to something outside me (that is, to something in another region of space from that in which I find myself), and similarly in order that I may be able to represent them as outside and alongside one another, and accordingly as not only different but as in different places, the representation of space must be presupposed. The representation of space cannot, therefore, be empirically obtained from the relations of outer appearance. On the contrary, this outer experience is itself possible at all only through that representation. 2. Space is a necessary a priori representation, which underlies all outer intuitions. We can never represent to ourselves the absence of space, though we can quite well think it as empty of objects. It must therefore be regarded as the condition of the possibility of appearances, and not as a determination dependent upon them. It is an a priori representation, which necessarily underlies outer appearances. * 3. The apodeictic certainty of all geometrical propositions and the possibility of their a priori construction is grounded in this a priori necessity of space. ....."

He is equally clear about the necessity of time as part of experience but he has no clear exposition of the (specious present) extended present:

P 74 "1. Time is not an empirical concept that has been derived from any experience. For neither coexistence nor succession would ever come within our perception, if the representation of time were not presupposed as underlying them a priori. Only on the presupposition of time can we represent to ourselves a number of things as existing at one and the same time (simultaneously) or at different times (successively). They are connected with the appearances only as effects accidentally added by the particular constitution of the sense organs. Accordingly, they are not a priori representations, but are grounded in sensation, and, indeed, in the case of taste, even upon feeling (pleasure and pain), as an effect of sensation. Further, no one can have a priori a representation of a colour or of any taste; whereas, since space concerns only the pure form of intuition, and therefore involves no sensation whatsoever, and nothing empirical, all kinds and determinations of space can and must be represented a priori, if concepts of figures and of their relations are to arise. Through space alone is it possible that things should be outer objects to us. ..2. 3.. 4.. 5..."

Kant has a model of experience as a succession of 3D instants, based on conventional 18th century thinking, allowing his reason to overcome his observation. He says of time that:

P 79 " It is nothing but the form of our inner intuition. If we take away from our inner intuition the peculiar condition of our sensibility, the concept of time likewise vanishes; it does not inhere in the objects, but merely in the subject which intuits them. I can indeed say that my representations follow one another; but this is only to say that we are conscious of them as in a time sequence, that is, in conformity with the form of inner sense. Time is not, therefore, something in itself, nor is it an objective determination inherent in things."

This analysis is strange because if uses the geometric term "form" but then uses the processing term "succession".

Gottfried Wilhelm Leibniz (1646-1716)

Leibniz is one of the first to notice that there is a problem with the proposition that computational machines could be conscious:

"One is obliged to admit that perception and what depends upon it is inexplicable on mechanical principles, that is, by figures and motions. In imagining that there is a machine whose construction would enable it to think, to sense, and to have perception, one could conceive it enlarged while retaining the same proportions, so that one could enter into it, just like into a windmill. Supposing this, one should, when visiting within it, find only parts pushing one another, and never anything by which to explain a perception. Thus it is in the simple substance, and not in the composite or in the machine, that one must look for perception." Monadology, 17.

Leibniz considered that the world was composed of "monads":

"1. The Monad, of which we shall here speak, is nothing but a simple substance, which enters into compounds. By 'simple' is meant 'without parts.' (Theod. 10.)

2. And there must be simple substances, since there are compounds; for a compound is nothing but a collection or aggregatum of simple things.

3. Now where there are no parts, there can be neither extension nor form [figure] nor divisibility. These Monads are the real atoms of nature and, in a word, the elements of things. " (Monadology 1714).

These monads are considered to be capable of perception through the meeting of things at a point:

"They cannot have shapes, because then they would have parts; and therefore one monad in itself, and at a moment, cannot be distinguished from another except by its internal qualities and actions; which can only be its *perceptions* (that is, the representations of the composite, or of what is external, in the simple), or its *appetitions* (its tending to move from one perception to another, that is), which are the principles of change. For the simplicity of a substance does not in any way rule out a multiplicity in the modifications which must exist together in one simple substance; and those modifications must consist in the variety of its relationships to things outside it - like the way in which in a *centre*, or a *point*, although it is completely simple, there are an infinity of angles formed which meet in it." (Principles of Nature and Grace 1714).

Leibniz also describes this in his "New System":

"It is only atoms of substance, that is to say real unities absolutely devoid of parts, that can be the sources of actions, and the absolute first principles of the composition of things, and as it were the ultimate elements in the analysis of substances <substantial things>. They might be called *metaphysical points*; they have *something of the nature of life* and a kind of *perception*, and *mathematical points* are their *point of view* for expressing the universe."(New System (11) 1695).

Having identified perception with metaphysical points Leibniz realises that there is a problem connecting the points with the world (cf: epiphenomenalism):

"Having decided these things, I thought I had reached port, but when I set myself to think about the union of the soul with the body I was as it were carried back into the open sea. For I could find no way of explaining how the body can make something pass over into the soul or vice versa, or how one created substance can communicate with another."(New System (12) 1695).

Leibniz devises a theory of "pre-established harmony" to overcome this epiphenomenalism. He discusses how two separate clocks could come to tell the same time and proposes that this could be due to mutual influence of one clock on the other ("the way of influence"), continual adjustment by a workman ("the way of assistance") or by making the clocks so well that they are always in agreement ("the way of pre-established agreement" or harmony). He considers each of these alternatives for harmonising the perceptions with the world and concludes that only the third is viable:

"Thus there remains only my theory, *the way of pre-established harmony*, set up by a contrivance of divine foreknowledge, which formed each of these substances from the outset in so perfect, so regular, and so exact a manner, that merely by following out its own laws, which were given to it when it was brought into being, each substance is nevertheless in harmony with the other, just as if there were a mutual influence between them, or as if in addition to his general concurrence God were continually operating upon them. (Third Explanation of the New System (5), 1696)."

This means that he must explain how perceptions involving the world take place:

"Because of the plenitude of the world everything is linked, and every body acts to a greater or lesser extent on every other body in proportion to distance, and is affected by it in return. It therefore follows that every monad is a living mirror, or a mirror endowed with internal activity, representing the universe in accordance with its own point of view, and as orderly as the universe itself. The perceptions of monads arise one out of another by the laws of appetite, or of the *final causes of good and evil* (which are prominent perceptions, orderly or disorderly), just as changes in bodies or in external phenomena arise one from another by the laws of *efficient causes*, of motion that is. Thus there is perfect *harmony* between the perceptions of the monad and the motions of bodies, pre-established from the outset, between the system of efficient causes and that of final causes. And it is that harmony that the agreement or physical union between the soul and body consists, without either of them being able to change the laws of the other." (Principles of Nature and Grace (3) 1714).

The "laws of appetite" are defined as:

"The action of the internal principle which brings about change, or the passage from one perception to another, can be called *appetition*. In fact appetite cannot always attain in its entirety the whole of the perception towards which it tends, but it always obtains some part of it, and attains new perceptions. Monadology 15.

Leibniz thought animals had souls but not minds:

"But *true* reasoning depends on necessary or eternal truths like those of logic, numbers, and geometry, which make indubitable connections between ideas, and conclusions which are inevitable. Animals in which such conclusions are never perceived are called *brutes*; but those which recognise such necessary truths are what are rightly called *rational animals* and their souls are called *minds*. (Principles of Nature and Grace (5) 1714).

Minds allow reflection and awareness:

"And it is by the knowledge of necessary truths, and by the abstractions they involve, that we are raised to *acts of reflection*, which make us aware of what we call *myself*, and make us think of this or that thing as in *ourselves*. And in this way, by thinking of ourselves, we think of being, of substance, of simples and composites, of the immaterial - and, by realising that what is limited in us is limitless in him, of God himself. And so these *acts of reflection* provide the principle objects of our reasonings." Monadology, 30.

George Berkeley (1685 - 1753)

A Treatise on the Principles of Human Knowledge. 1710

http://darkwing.uoregon.edu/~rbear/berkeley.html#treatise

Berkeley introduces the Principles of Human Knowledge with a diatribe against abstract ideas. He uses the abstract ideas of animals as an example:

"Introduction. 9.....The constituent parts of the abstract idea of animal are body, life, sense, and spontaneous motion. By body is meant body without any particular shape or figure, there being no one shape or figure common to all animals, without covering, either of hair, or feathers, or scales, &c., nor yet naked: hair, feathers, scales, and nakedness being the distinguishing properties of particular animals, and for that reason left out of the abstract idea. Upon the same account the spontaneous motion must be neither walking, nor flying, nor creeping; it is nevertheless a motion, but what that motion is it is not easy to conceive.

He then declares that such abstractions cannot be imagined. He emphasises that ideas are "represented to myself" and have shape and colour:

"Introduction. 10. Whether others have this wonderful faculty of abstracting their ideas, they best can tell: for myself, I find indeed I have a faculty of imagining, or representing to myself, the ideas of those particular things I have perceived, and of variously compounding and dividing them. I can imagine a man with two heads, or the upper parts of a man joined to the body of a horse. I can consider the hand, the eye, the nose, each by itself abstracted or separated from the rest of the body. But then whatever hand or eye I imagine, it must have some particular shape and colour. Likewise the idea of man that I frame to myself must be either of a white, or a black, or a tawny, a straight, or a crooked, a tall, or a low, or a middle-sized man. I cannot by any effort of thought conceive the abstract idea above described. And it is equally impossible for me to form the abstract idea of motion distinct from the body moving, and which is neither swift nor slow, curvilinear nor rectilinear; and the like may be said of all other abstract general ideas whatsoever."

This concept of ideas as extended things, or representations, is typical of the usage amongst philosophers in the 17th and 18th century and can cause confusion in modern readers. Berkeley considers that words that are used to describe classes of things in the abstract can only be conceived as particular cases:

"Introduction. 15... Thus, when I demonstrate any proposition concerning triangles, it is to be supposed that I have in view the universal idea of a triangle; which ought not to be understood as if I could frame an idea of a triangle which was neither equilateral, nor scalenon, nor equicrural; but only that the particular triangle I consider, whether of this or that sort it matters not, doth equally stand for and represent all rectilinear triangles whatsoever, and is in that sense universal. All which seems very plain and not to include any difficulty in it.

Intriguingly, he considers that language is used to directly excite emotions as well as to communicate ideas:

"Introduction. 20. ... I entreat the reader to reflect with himself, and see if it doth not often happen, either in hearing or reading a discourse, that the passions of fear, love, hatred, admiration, disdain, and the like, arise immediately in his mind upon the perception of certain words, without any ideas coming between.

Berkeley considers that extension is a quality of mind:

"11. Again, great and small, swift and slow, are allowed to exist nowhere without the mind, being entirely relative, and changing as the frame or position of the organs of sense varies. The extension

therefore which exists without the mind is neither great nor small, the motion neither swift nor slow, that is, they are nothing at all. But, say you, they are extension in general, and motion in general: thus we see how much the tenet of extended movable substances existing without the mind depends on the strange doctrine of abstract ideas."

He notes that the rate at which things pass may be related to the mind:

"14..... Is it not as reasonable to say that motion is not without the mind, since if the succession of ideas in the mind become swifter, the motion, it is acknowledged, shall appear slower without any alteration in any external object?

Berkeley raises the issue of whether objects exist without being perceived. He bases his argument on the concept of perception being the perceiving of "our own ideas or sensations":

"4. It is indeed an opinion strangely prevailing amongst men, that houses, mountains, rivers, and in a word all sensible objects, have an existence, natural or real, distinct from their being perceived by the understanding. But, with how great an assurance and acquiescence soever this principle may be entertained in the world, yet whoever shall find in his heart to call it in question may, if I mistake not, perceive it to involve a manifest contradiction. For, what are the fore-mentioned objects but the things we perceive by sense? and what do we perceive besides our own ideas or sensations? and is it not plainly repugnant that any one of these, or any combination of them, should exist unperceived?"

He further explains this concept in terms of some Eternal Spirit allowing continued existence. Berkeley is clear that the contents of the mind have "colour, figure, motion, smell, taste etc.":

"7. From what has been said it follows there is not any other Substance than Spirit, or that which perceives. But, for the fuller proof of this point, let it be considered the sensible qualities are colour, figure, motion, smell, taste, etc., i.e. the ideas perceived by sense. Now, for an idea to exist in an unperceiving thing is a manifest contradiction, for to have an idea is all one as to perceive; that therefore wherein colour, figure, and the like qualities exist must perceive them; hence it is clear there can be no unthinking substance or substratum of those ideas."

He elaborates the concept that there is no unthinking substance or substratum for ideas and all is mind:

"18. But, though it were possible that solid, figured, movable substances may exist without the mind, corresponding to the ideas we have of bodies, yet how is it possible for us to know this? Either we must know it by sense or by reason. As for our senses, by them we have the knowledge only of our sensations, ideas, or those things that are immediately perceived by sense, call them what you will: but they do not inform us that things exist without the mind, or unperceived, like to those which are perceived. This the materialists themselves acknowledge. It remains therefore that if we have any knowledge at all of external things, it must be by reason, inferring their existence from what is immediately perceived by sense. But what reason can induce us to believe the existence of bodies without the mind, from what we perceive, since the very patrons of Matter themselves do not pretend there is any necessary connexion betwixt them and our ideas? I say it is granted on all hands (and what happens in dreams, phrensies, and the like, puts it beyond dispute) that it is possible we might be affected with all the ideas we have now, though there were no bodies existing without resembling them. Hence, it is evident the supposition of external bodies is not necessary for the producing our ideas; since it is granted they are produced sometimes, and might possibly be produced always in the same order, we see them in at present, without their concurrence. "

and stresses that there is no apparent connection between mind and the proposed material substrate of ideas:

"19. But, though we might possibly have all our sensations without them, yet perhaps it may be thought easier to conceive and explain the manner of their production, by supposing external bodies in their likeness rather than otherwise; and so it might be at least probable there are such things as

bodies that excite their ideas in our minds. But neither can this be said; for, though we give the materialists their external bodies, they by their own confession are never the nearer knowing how our ideas are produced; since they own themselves unable to comprehend in what manner body can act upon spirit, or how it is possible it should imprint any idea in the mind. .....

Berkeley makes a crucial observation, that had also been noticed by Descartes, that ideas are passive:

"25. All our ideas, sensations, notions, or the things which we perceive, by whatsoever names they may be distinguished, are visibly inactive- there is nothing of power or agency included in them. So that one idea or object of thought cannot produce or make any alteration in another. To be satisfied of the truth of this, there is nothing else requisite but a bare observation of our ideas. For, since they and every part of them exist only in the mind, it follows that there is nothing in them but what is perceived: but whoever shall attend to his ideas, whether of sense or reflexion, will not perceive in them any power or activity; there is, therefore, no such thing contained in them. A little attention will discover to us that the very being of an idea implies passiveness and inertness in it, insomuch that it is impossible for an idea to do anything, or, strictly speaking, to be the cause of anything: neither can it be the resemblance or pattern of any active being, as is evident from sect. 8. Whence it plainly follows that extension, figure, and motion cannot be the cause of our sensations. To say, therefore, that these are the effects of powers resulting from the configuration, number, motion, and size of corpuscles, must certainly be false.

He considers that "the cause of ideas is an incorporeal active substance or Spirit (26)".

He summarises the concept of an Eternal Spirit that governs real things and a representational mind that copies the form of the world as follows:

"33. The ideas imprinted on the Senses by the Author of nature are called real things; and those excited in the imagination being less regular, vivid, and constant, are more properly termed ideas, or images of things, which they copy and represent. But then our sensations, be they never so vivid and distinct, are nevertheless ideas, that is, they exist in the mind, or are perceived by it, as truly as the ideas of its own framing. The ideas of Sense are allowed to have more reality in them, that is, to be more strong, orderly, and coherent than the creatures of the mind; but this is no argument that they exist without the mind. They are also less dependent on the spirit, or thinking substance which perceives them, in that they are excited by the will of another and more powerful spirit; yet still they are ideas, and certainly no idea, whether faint or strong, can exist otherwise than in a mind perceiving it.

Berkeley considers that the concept of distance is a concept in the mind and also that dreams can be compared directly with sensations:

"42. Thirdly, it will be objected that we see things actually without or at distance from us, and which consequently do not exist in the mind; it being absurd that those things which are seen at the distance of several miles should be as near to us as our own thoughts. In answer to this, I desire it may be considered that in a dream we do oft perceive things as existing at a great distance off, and yet for all that, those things are acknowledged to have their existence only in the mind."

He considers that ideas can be extended without the mind being extended:

"49. Fifthly, it may perhaps be objected that if extension and figure exist only in the mind, it follows that the mind is extended and figured; since extension is a mode or attribute which (to speak with the schools) is predicated of the subject in which it exists. I answer, those qualities are in the mind only as they are perceived by it- that is, not by way of mode or attribute, but only by way of idea; and it no more follows the soul or mind is extended, because extension exists in it alone, than it does that it is red or blue, because those colours are on all hands acknowledged to exist in it, and nowhere else. As to what philosophers say of subject and mode, that seems very groundless and unintelligible. For instance, in this proposition "a die is hard, extended, and square," they will have it that the word die denotes a subject or substance, distinct from the hardness, extension, and figure

which are predicated of it, and in which they exist. This I cannot comprehend: to me a die seems to be nothing distinct from those things which are termed its modes or accidents. And, to say a die is hard, extended, and square is not to attribute those qualities to a subject distinct from and supporting them, but only an explication of the meaning of the word die."

Berkeley proposes that time is related to the succession of ideas:

"98. For my own part, whenever I attempt to frame a simple idea of time, abstracted from the succession of ideas in my mind, which flows uniformly and is participated by all beings, I am lost and embrangled in inextricable difficulties. I have no notion of it at all, only I hear others say it is infinitely divisible, and speak of it in such a manner as leads me to entertain odd thoughts of my existence; since that doctrine lays one under an absolute necessity of thinking, either that he passes away innumerable ages without a thought, or else that he is annihilated every moment of his life, both which seem equally absurd. Time therefore being nothing, abstracted from the sucession of ideas in our minds, it follows that the duration of any finite spirit must be estimated by the number of ideas or actions succeeding each other in that same spirit or mind. Hence, it is a plain consequence that the soul always thinks; and in truth whoever shall go about to divide in his thoughts, or abstract the existence of a spirit from its cogitation, will, I believe, find it no easy task.

"99. So likewise when we attempt to abstract extension and motion from all other qualities, and consider them by themselves, we presently lose sight of them, and run into great extravagances. All which depend on a twofold abstraction; first, it is supposed that extension, for example, may be abstracted from all other sensible qualities; and secondly, that the entity of extension may be abstracted from its being perceived. But, whoever shall reflect, and take care to understand what he says, will, if I mistake not, acknowledge that all sensible qualities are alike sensations and alike real; that where the extension is, there is the colour, too, i.e., in his mind, and that their archetypes can exist only in some other mind; and that the objects of sense are nothing but those sensations combined, blended, or (if one may so speak) concreted together; none of all which can be supposed to exist unperceived."

He regards "spirit" as something separate from ideas and attempts to answer the charge that as spirit is not an idea it cannot be known:

"139. But it will be objected that, if there is no idea signified by the terms soul, spirit, and substance, they are wholly insignificant, or have no meaning in them. I answer, those words do mean or signify a real thing, which is neither an idea nor like an idea, but that which perceives ideas, and wills, and reasons about them. ....

Thomas Reid (1710-1796)

Thomas Reid is generally regarded as the founder of Direct Realism. Reid was a Presbyterian minister for the living of Newmachar near Aberdeen from 1737. He is explicit about the 'directness' of his realism:

"It is therefore acknowledged by this philosopher to be a natural instinct or prepossession, a universal and primary opinion of all men, a primary instinct of nature, that the objects which we immediately perceive by our senses are not images in our minds, but external objects, and that their existence is independent of us and our perception. (Thomas Reid Essays, 14)"

In common with Descartes and Malebranche, Reid considers that the mind itself is an unextended thing:

".. I take it for granted, upon the testimony of common sense, that my mind is a substance-that is, a permanent subject of thought; and my reason convinces me that it is an unextended and invisible

substance; and hence I infer that there cannot be in it anything that resembles extension (Inquiry)".

Reid is also anxious to equate the unextended mind with the soul:

"The soul, without being present to the images of the things perceived, could not possibly perceive them. A living substance can only there perceive, where it is present, either to the things themselves, (as the omnipresent God is to the whole universe,) or to the images of things, as the soul is in its proper sensorium."

Reid's Direct Realism is therefore the idea that the physical objects in the world are in some way presented directly to a soul. This approach is known as "Natural Dualism".

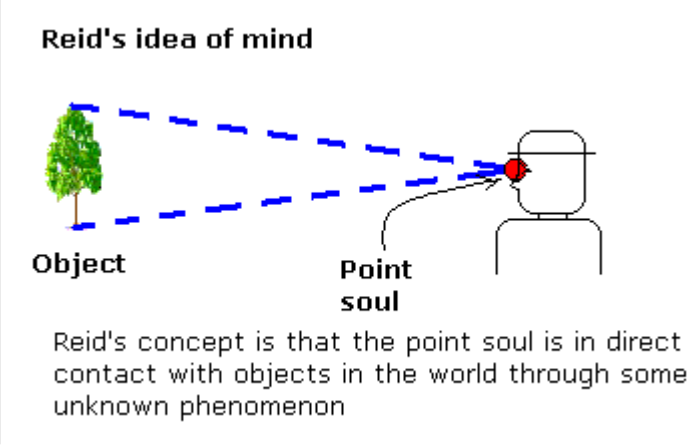
Reid's views show his knowledge of Aristotle's ideas:

"When we perceive an object by our senses, there is, first, some impression made by the object upon the organ of sense, either immediately, or by means of some medium. By this, an impression is made upon the brain, in consequence of which we feel some sensation. " (Reid 1785)

He differs from Aristotle because he believes that the content of phenomenal consciousness is things in themselves, not signals derived from things in the brain. However, he has no idea how such a phenomenon could occur:

"How a sensation should instantly make us conceive and believe the existence of an external thing altogether unlike it, I do not pretend to know; and when I say that the one suggests the other, I mean not to explain the manner of their connection, but to express a fact, which everyone may be conscious of namely, that, by a law of our nature, such a conception and belief constantly and immediately follow the sensation." (Reid 1764).

Reid's idea of mind is almost impossible to illustrate because it lacks sufficient physical definition. It is like naive realism but without any communication by light between object and observer. Reid was largely ignored until the rise of modern Direct Realism.



Reading between the lines, it seems that Reid is voicing the ancient intuition that the observer and the content of an observation are directly connected in some way. As will be seen later, this intuition cannot distinguish between a direct connection with the world itself and a direct connection with signals from the world beyond the body that are formed into a virtual reality in the brain.

References

- Descartes, R. (1628). Rules For The Direction of The Mind. http://faculty.uccb.ns.ca/philosophy/kbryson/rulesfor.htm
- Descartes, R. (1637). DISCOURSE ON THE METHOD OF RIGHTLY CONDUCTING THE REASON, AND SEEKING TRUTH IN THE SCIENCES. http://www.gutenberg.org/dirs/etext93/dcart10.txt
- Descartes, R. (1641). Meditations on First Philosophy. http://oregonstate.edu/instruct/phl302/texts/descartes/meditations/Meditation1.html
- Descartes, R. (1664) "Treatise on Man". Translated by John Cottingham, et al. The Philosophical Writings of Descartes, Vol. 1 (Cambridge: Cambridge University Press, 1985) 99-108. http://lrc.csun.edu/~battias/454/descartes/
- Kant, I. (1781) Critique of Pure Reason. Trans. Norman Kemp Smith with preface by Howard Caygill. Pub: Palgrave Macmillan. http://www.hkbu.edu.hk/~ppp/cpr/toc.html
- Locke, J. (1689). An Essay Concerning Human Understanding http://www.arts.cuhk.edu.hk/Philosophy/Locke/echu/
- Russell, B. (1945). A History of Western Philosophy. New York: Simon and Schuster.
- Reid, T. (1785). Essays on the Intellectual Powers of Man. Edited by Brookes, Derek. Edinburgh: Edinburgh University Press, 2002.
- Reid, T. (1764). An Inquiry into the Human Mind on the Principles of Common Sense. Edited by Brookes, Derek. Edinburgh: Edinburgh University Press, 1997.

Further Reading

See: THOMAS REID at LoveToKnow. http://54.1911encyclopedia.org/R/RE/REID_THOMAS.htm

- Consciousness in Descartes

Nineteenth and twentieth century philosophy of consciousness

The nineteenth and twentieth centuries witnessed a confident use of nineteenth century scientific ideas amongst philosophers of mind and a few philosophers such as Whitehead were also coming to terms with modern science.

ER Clay

ER Clay deserves a mention in the catalogue of important nineteenth century philosophers of consciousness for the quotation from his work given in William James' classic text *The Principles of Psychology*:

"The relation of experience to time has not been profoundly studied. Its objects are given as being of the present, but the part of time referred to by the datum is a very different thing from the conterminous of the past and future which philosophy denotes by the name Present. The present to which the datum refers is really a part of the past -- a recent past -- delusively given as being a time that intervenes between the past and the future. Let it be named the specious present, and let the past, that is given as being the past, be known as the obvious past. All the notes of a bar of a song seem to the listener to be contained in the present. All the changes of place of a meteor seem to the beholder to be contained in the present. At the instant of the termination of such series, no part of the time measured by them seems to be a past. Time, then, considered relatively to human apprehension, consists of four parts, viz., the obvious past, the specious present, the real present, and the future. Omitting the specious present, it consists of three . . . nonentities -- the past, which does not exist, the future, which does not exist, and their conterminous, the present; the faculty from which it proceeds lies to us in the fiction of the specious present."

Clay provides an eloquent description of the extended, or *specious*, present, mentioning both the way that consciousness seems to occupy a duration of time and the way that events within conscious experience have their own durations so that they snap out of existence when they end. This description in itself allows us to see how McTaggart's "A Series" might be constructed from the overlapping extended present's of events.

Clay's use of the pejorative term "specious" for the way that experience has a duration was necessary in the nineteenth century but now we know that it was the nineteenth century idea of physical time that was specious. A neutral term for experience laid out in time might be the "extended present".

Alfred North Whitehead

The Concept of Nature. Cambridge: Cambridge University Press (1920): 49-73.

Many twentieth century philosophers have taken the nineteenth century idea of space and time as the framework within which their descriptions of experience are elaborated. Whitehead was a mathematician and philosopher who understood the limitations of this framework and pointed out that our failure to understand and overcome these limitations was probably at the root of our failure to understand consciousness. He traces the problem to the nineteenth century view of time and space and rails against materialists who elevate nineteenth century scientific doctrine above observational and scientific reality.

He also believed that mind and nature are part of the same phenomena:

"What I am essentially protesting against is the bifurcation of nature into two systems of reality, which, in so far as they are real, are real in different senses. One reality would be the entities such as electrons which are the study of speculative physics. This would be the reality which is there for knowledge; although on this theory it is never known. For what is known is the other sort of reality, which is the byplay of the mind. Thus there would be two natures, one is the conjecture and the other is the dream.

"Another way of phrasing this theory which I am arguing against is to bifurcate nature into two divisions, (31) namely into the nature apprehended in awareness and the nature which is the cause of awareness. The nature which is the fact apprehended in awareness holds within it the greenness of the trees, the song of the birds, the warmth of the sun, the hardness of the chairs, and the feel of the velvet. The nature which is the cause of awareness is the conjectured system of molecules and electrons which so affects the mind as to produce the awareness of apparent nature. The meeting point of these two natures is the mind, the causal nature being influent and the apparent nature being effluent."

He argued that science is about the relations between things:

"The understanding which is sought by science is an understanding of relations within nature. "

Whitehead was aware of the way that the simultaneity of events is of crucial importance to phenomenal experience:

"The general fact is the whole simultaneous occurrence of nature which is now for sense-awareness. This general fact is what I have called the discernible. But in future I will call it a 'duration,' meaning thereby a certain whole of nature which is limited only by the property of being a simultaneity. Further in obedience to the principle of comprising within nature the whole terminus of sense-awareness, simultaneity must not be conceived as an irrelevant mental concept imposed upon nature. Our sense-awareness posits for immediate discernment a certain whole, here called a 'duration'; thus a duration is a definite natural entity. A duration is discriminated as a complex of partial events, and the natural entities which are components of this complex are thereby said to be 'simultaneous with this duration.' Also in a derivative sense they are simultaneous with each other in respect to this duration. Thus simultaneity is a definite natural relation. The word' duration' is perhaps unfortunate in so far as it suggests a mere abstract stretch of time. This is not what I mean. A duration is a concrete slab of nature limited by simultaneity which is an essential factor disclosed in sense-awareness."

Whitehead also stresses the role of the extended, or 'specious', present in sense awareness:

"It is important to distinguish simultaneity from instantaneousness. I lay no stress on the mere current usage of the two terms. There are two concepts which I want to distinguish, and one I call simultaneity and the other instantaneousness. I hope that the words are judiciously chosen; but it really does not matter so long as I succeed in explaining my meaning. Simultaneity is the property of a group of natural elements which in some sense are components of a duration. A duration can be all nature present as the immediate fact posited by sense-awareness. A duration retains within itself the passage of nature. There are within it antecedents and consequents which are also durations which may be the complete specious presents of quicker consciousnesses. In other words a duration retains temporal thickness. Any concept of all nature as immediately known is always a concept of some duration though it may be enlarged in its temporal thickness beyond the possible specious present of any being known to us as existing within nature. Thus simultaneity is an ultimate factor in nature, immediate for sense-awareness.

So a set of events that are extended in time constitutes conscious experience. He then defines continuity in terms of overlapping durations:

"The continuity of nature arises from extension. Every event extends over other events, and every event is extended over by other events. Thus in the special case of durations which are now the only events directly under consideration, every duration is part of other durations; and every duration has

other durations which are parts of it."

That experience exists as whole durations that overlap means that the overlapping durations can be considered to be composed of moments or instants and these can be assigned to a series which we call 'time':

"Such an ordered series of moments is what we mean by time defined as a series. Each element of the series exhibits an instantaneous; state of nature, Evidently this serial time is the result of an intellectual process of (65) abstraction."

Processes can occur within a duration of sense awareness so things can change within the extended present of a conscious interval.

"Sense-awareness and thought are themselves processes as well as their termini in nature."

So Whitehead's durations of sense awareness both contain processes and are phenomena in their own right. A movement can be both a succession of changes of position and a quality of motion over the whole duration that contains it.

One disturbing feature of his analysis is that he does not mention the way that durations are attached to events; Clay states that the extension in time of an event disappears when the event ceases.

Whitehead recognised a possible problem with Aristotle's system of categories. Whitehead suggested substituting his category of relation for Aristotle's category of substance, giving us a different conception of being. This allowed Whitehead to postulate a conception of consciousness that avoided an infinite regress and the need for accounting for awareness as a product of our physical senses. Thus, a diiferent approach was recognized, one that rivaled the post-Cartesian approach to understanding the world. It is Whitehead's encounter with Aristotle that allowed him to formulate a different possible solution to the problem of consciousness and one that falls squarely in the domain of panpsychism.

Edmund Husserl

Husserl accepts the materialist paradigm and has been influential in Marxist and post-Marxist philosophy. Husserl writes in a style that presents a multitude of views, many of which are opposed to each other. He is also rather obscure when concepts become difficult, an example of this postmodern penchant for confusion is given below:

"The genuine intentional synthesis is discovered in the synthesis of several acts into one act, such that, in a unique manner of binding one meaning to another, there emerges not merely a whole, an amalgam whose parts are meanings, but rather a single meaning in which these meanings themselves are contained, but in a meaningful way. With this the problems of correlation, too, already announce themselves; and thus, in fact, this work contains the first, though of course very imperfect, beginnings of "phenomenology."" (Husserl 1937).

Husserl seems to be largely a Humean in the sense that he gives precedence to mental experience as the only thing that may be known directly and hence certainly. He regards the components of experience as part of consciousness, so the intention to move, the movement and the sensation of movement are bound or 'bracketed' together into a single meaning.

"In my perceptual field I find myself holding sway as ego through my organs and generally through everything belonging to me as an ego in my ego-acts and faculties. However, though the objects of the life-world, if they are to show their very own being, necessarily show themselves as physical bodies, this does not mean that they show themselves only in this way; and [similarly] we, though we are related through the living body to all objects which exist for us, are not related to them solely as a living body. Thus if it is a question of objects in the perceptual field, we are perceptually also in the field; and the same is true, in modification, of every intuitive field, and even of every nonintuitive one, since we are obviously capable of "representing" to ourselves everything which is non-intuitively before us (though we are sometimes temporally limited in this). [Being related] "through the living body" clearly does not mean merely [being related] "as a physical body"; rather, the expression refers to the kinesthetic, to functioning as an ego in this peculiar way, primarily through seeing, hearing, etc.; and of course other modes of the ego belong to this (for example, lifting, carrying, pushing, and the like). "

It should be noted that Husserl believes we *perform* acts of perception and that we should refrain from judgement about where the things in perception are located or their nature. This suspenson of judgement is called *epoche* and derives from ancient Greek skepticism.

Despite an affection for long sentences with dubious meanings Husserl seems to share Locke's view that experience is extended in time. He is obscure about whether he believes consciousness itself is a process that initiates action. Husserl uses a linguistic argument to justify the idea of consciousness as a form of action:

"2. Whatever becomes accessible to us through reflection has a noteworthy universal character: that of being consciousness of something, of having something as an object of consciousness, or correlatively, to be aware of it we are speaking here of intentionality. This is the essential character of mental life in the full sense of the word, and is thus simply inseparable from it. It is, for example, inseparable from the perceiving that reflection reveals to us, that it is of this or that; just as the process of remembering is, in itself, remembering or recalling of this or that; just as thinking is thinking of this or that thought, fearing is of something, love is of something; and so on. We can also bring in here the language we use in speaking of appearing or having something appear."(Husserl 1928)

Intentionality is mentioned but not described. Intentionality is a process and Husserl seems to be suggesting that consciousness is a process but he does not describe any consciousness of the transformation that is this process. He simply assumes, as a cornerstone of his approach, that consciousness is a process:

"5. The Purely Mental in Experience of the Self and of Community. The All-Embracing Description of Intentional Processes." (Husserl 1928).

then, not surprisingly, fails to find any processes within it and changes his view of consciousness to that of observation:

"... But I <must> immediately add that the universality of the phenomenological epoche as practiced by the phenomenologist from the very beginning the universality in which he or she becomes the mere impartial observer of the totality of his conscious life-process brings about not only a thematic purification of the individual processes of consciousness and thereby discloses its noematic components;" (Husserl 1928)

He calls the contents of perception the perceptual *noema*. Husserl seems to be aware of the problem of the extended present:

"How can we account for the fact that a presently occurring experience in one's consciousness called "recollection" makes us conscious of a not-present event and indeed makes us aware of it as past? And how is it that in the "remembered" moment, that sense can be included in an evidential

way with the sense: "have earlier perceived"? How are we to understand the fact that a perceptual, that is to say, bodily characterized present can at the same time contain a co-presence with the sense of a perceivability that goes beyond the <immediate> perceivedness? How are we to understand the fact that the actual perceptual present as a totality does not close out the world but rather always carries within itself the sense of an infinite plus ultra <more beyond>?"(Husserl 1928)

But is vague about whether mental time is a continuum or has three components of remembered past, present and some sort of intuition of the future. His rejection of the possibility of describing the mind through the spatio-temporal models of the physical sciences limits his interpretation of mental space and time.

Husserl, E. (1928) The Amsterdam Lectures. PSYCHOLOGICAL AND TRANSCENDENTAL PHENOMENOLOGY AND THE CONFRONTATION WITH HEIDEGGER (1927-1931). edited and translated by Thomas Sheehan and Richard E. Palmer. http://www.stanford.edu/dept/relstud/faculty/sheehan.bak/TSbookcontents.html

Husserl, E. (1937). The Crisis of European Sciences and Transcendental Phenomenology. An Introduction to Phenomenology. (The Crisis of European Sciences and Transcendental Phenomenology (1954) publ. Northwestern University Press, Evanston, 1970. Sections 22 - 25 and 57 - 68, 53 pages in all.) http://www.marxists.org/reference/subject/philosophy/works/ge/husserl.htm

Daniel Clement Dennett (1942 -)

Dennett is well known for his "Multiple Drafts Model" of consciousness. The Multiple Drafts Theory or Model of Consciousness is a theory of consciousness based upon the proposal that the brain acts as an information processor. The Theory is described in depth in the book Consciousness Explained, written by Dennett in 1991. It proposes a form of strong AI.

Dennet describes his theory (CE p117) as **operationalist**, as Dennett says: "There is no reality of conscious experience independent of the effects of various vehicles of content on subsequent action (and hence, of course, on memory)." (Not to be confused with 'instrumentalism').

Dennett's starting point in the development of the Multiple Drafts theory is a description of the phi illusion. In this experiment two different coloured lights, with an angular separation of a few degrees at the eye, are flashed in succession. If the interval between the flashes is less than a second or so the first light that is flashed appears to move across to the position of the second light. Furthermore the light seems to change colour as it moves across the visual field. A green light will appear to turn red as it seems to move across to the position of a red light. Dennett asks how we could see the light change colour before the second light is observed.

An example of the phi illusion in the format described by Dennett is shown here: phi illusion (use the 'test' option to select the simple phi demonstration).

Dennett explains the change of colour of the light in terms of either Orwellian or Stalinesque hypotheses. In the Orwellian hypothesis the subject develops a narrative about the movement of the lights after the event. In the Stalinesque hypothesis the subject's brain would have a delay in which the movement of the green light towards the red light could be modelled after the sensory information from the red light had been received. He then says that it does not matter which hypothesis applies because: "the Multiple Drafts model goes on to claim that the brain does not bother 'constructing' any representations that go to the trouble of 'filling in' the blanks. That would be a waste of time and (shall we say?) paint. The judgement is already in so we can get on with other tasks!"

According to the Multiple Drafts theory there are a variety of sensory inputs from a given event and also a variety of interpretations of these inputs. The sensory inputs arrive in the brain and are interpreted at different times so a given event can give rise to a succession of discriminations. As soon as each discrimination is accomplished it becomes available for eliciting a behaviour. A wide range of behaviours may occur ranging from reactions to the event such as running away to descriptions of the experience of the event etc.

At different times after the event a person is able to relate different stories of what happened depending upon the extent to which the event has been analysed. Dennett compares this with a 'Cartesian Theatre' model of consciousness in which events suddenly appear on some sort of mental screen and then disappear as quickly. He provides numerous examples to show that events are analysed over a period of time rather than instantaneously.

Although Multiple Drafts is described as a model or theory of consciousness that differs from other models, Dennett points out that even Descartes was aware that reactions to an event could occur over a period of time with reflexes occurring first and judgements later. What makes Multiple Drafts different is that Dennett, in different sections of Consciousness Explained, either denies that normal conscious experiences actually occur or describes these as emerging in some unspecified way from the sheer complexity of information processing in the brain. His emergentism is clear when he defends the Multiple Drafts Model from Searle's chinese room argument by saying of the critics: They just can't imagine how understanding could be a property that emerges from lots of distributed quasi-understanding in a large system (CE p439).

As an example of denial of conscious experience Dennett denies that there is any internal experience of colour, instead he says that qualia in general are "mechanically accomplished dispositions to react". This view originates in Dennett's belief in the method of heterophenomenology in which narrative is thought to be the most crucial tool for investigating consciouness. However, Dennett does not deny conscious experience (see below).

The origin of this operationalist appoach can be seen in Dennett's immediately earlier work. Dennett (1988) redefines consciousness in terms of access consciousness alone, he argues that "Everything real has properties, and since I don't deny the reality of conscious experience, I grant that conscious experience has properties". Having related all consciousness to properties he then declares that these properties are actually judgements of properties. He considers judgements of the properties of consciousness to be identical to the properties themselves. He writes:

"The infallibilist line on qualia treats them as properties of one's experience one cannot in principle

misdiscover, and this is a mysterious doctrine (at least as mysterious as papal infal libility) unless we shift the emphasis a little and treat qualia as logical constructs out of subjects' qualia-judgments: a subject's experience has the quale F if and only if the subject judges his experience to have quale F."

Having identified "properties" with "judgement of properties" he can then show that the judgements are insubstantial, hence the properties are insubstantial and hence the qualia are insubstantial or even non-existent. Dennett concludes that qualia can be rejected as non-existent:

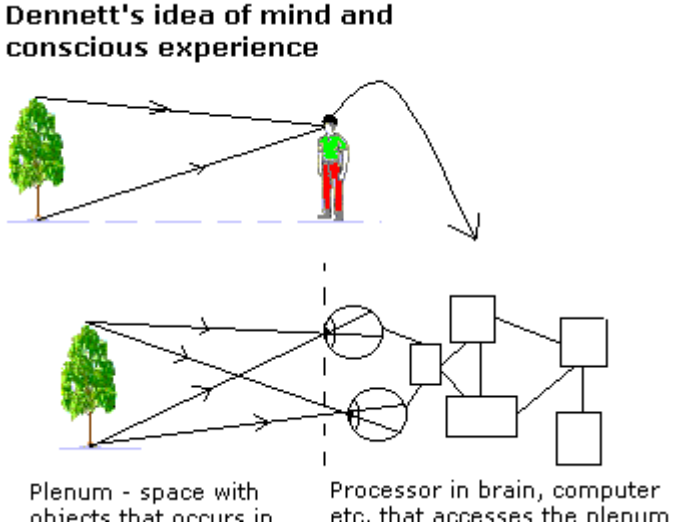
"So when we look one last time at our original characterization of qualia, as ineffable, intrinsic, private, directly apprehensible properties of experience, we find that there is nothing to fill the bill. In their place are relatively or practically ineffable public properties we can refer to indirectly via reference to our private property-detectors-- private only in the sense of idiosyncratic. And insofar as we wish to cling to our subjective authority about the occurrence within us of states of certain types or with certain properties, we can have some authority--not infallibility or incorrigibility, but something better than sheer guessing--but only if we restrict ourselves to relational, extrinsic properties like the power of certain internal states of ours to provoke acts of apparent re- identification. So contrary to what seems obvious at first blush, there simply are no qualia at all. " (Dennett 1988)

This identification of qualia with judgements rather than experience is the key to the Multiple Drafts Model, once accepted there is only a need to explain behaviour rather than personal experience itself.

The origin of this identification of qualia with judgements can be seen in *Consciousness Explained* p407-408. Dennett considers the experiences of someone looking at the world, and describes his idea of the relationship between conscious experience, mind and representation:

"It seemed to him, according to the text, as if his mind - his visual field - were filled with intricate details of gold-green buds and wiggling branches, but although this is how it seemed this was an illusion. No such "plenum" ever came into his mind; the plenum remained out in the world where it didn't have to be represented, but could just be. When we marvel, in those moments of heightened self-consciousness, at the glorious richness of our conscious experience, the richness we marvel at is actually the richness of the world outside, in all its ravishing detail. It does not "enter" our conscious minds, but is simply available"

For Dennett minds have no "plenum", no space with objects in it, the plenum is things outside the mind. Dennett considers mind to be processes. In his imaginary dialogue with 'Otto' in *Consciousness Explained* Dennett has Otto say "Are you denying then that consciousness is a plenum?" to which he replies "Yes indeed. That's part of whatI am denying. Consciousness is gappy and sparse, and doesn't contain half of what people think is there!". (CE p366). Unfortunately Dennett's assertion is difficult to understand because even half a plenum is a plenum, perhaps his remarks given above that 'conscious experience' ıtion. More than one thing at an num would seem to be equivalent



Dennett makes a sharp distinction between information in the world and information in the brain. The information in the world seems to be allowed to be a plenum that can enter conscious experience but ceases to be a plenum in the mind. In contrast, according to Dennett the information in the brain is a "logical space":

"So we do have a way of making sense of the idea of phenomenal space - as a logical space. This is a space into which or in which nothing is literally projected; its properties are simply constituted by the beliefs of the (heterophenomenological) subject."

Although how a "logical space" differs from a real space if it contains several things at an instant is not explained and how this "logical space" appears like phenomenal space at each instant is also not covered.

Dennett also attacks "Cartesian materialism" which he defines very precisely as the idea that there is a Cartesian theatre in the brain:

Lets call the idea of such a centered locus in the brain *Cartesian materialism*, since its the view you arrive at when you discard Descarte's dualism but fail to discard the imagery of a central (but material) Theater where "it all comes together". The pineal gland would be one candidate for such a Cartesian Theater, but there are others that have been suggested - the anterior cingulate, the reticular formation, various places in the frontal lobes. Cartesian materialism is the view that there is a crucial finish line or boundary somewhere in the brain, marking a place where the order of arrival equals the order of "presentation" in experience because what *happens* there is what you are conscious of."(CE p107)

It seems that Dennett is unaware of earlier uses of the term "Cartesian materialism" meaning the concept that the mind is in the brain and co-opts the term for his own use.

Dennett(1998) describes consciousness as distributed in time and space: "Consciousness doesn't have to happen at an instant; it is much better to think of it as distributed in both space and time." but, unlike Descartes, Broad or Whitehead uses an early materialist conception of time to describe it.

References

Daniel C Dennett. (1988). Quining Qualia. in A. Marcel and E. Bisiach, eds, Consciousness in Modern Science, Oxford University Press 1988. Reprinted in W. Lycan, ed., Mind and Cognition: A Reader, MIT Press, 1990, A. Goldman, ed. Readings in Philosophy and Cognitive Science, MIT Press, 1993. http://ase.tufts.edu/cogstud/papers/quinqual.htm

Dennett, D. C., 1998, The Myth of Double Transduction in S. R. Hameroff , A. W. Kaszniak, and A. C. Scott, eds., Toward a Science of Consciousness, II , Cambridge, MA: MIT Press/A Bradford Book, pp97-107.

Daniel C Dennett. (1991). Consciousness Explained. Little, Brown & Co. USA. Available as a Penguin Book.

Dennett, D. and Kinsbourne, M. (1992) Time and the Observer: the Where and When of Consciousness in the Brain. (1992) Behavioral and Brain Sciences, 15, 183-247, 1992. Reprinted in The Philosopher's Annual, Grim, Mar and Williams, eds., vol. XV-1992, 1994, pp. 23-68; Noel Sheehy and Tony Chapman, eds., Cognitive Science, Vol. I, Elgar, 1995, pp.210-274.http://ase.tufts.edu/cogstud/papers/time&obs.htm

Ned Block (1942-)

Ned Block is in the NYU Department of Philosophy.

Two types of consciousness

According to Block[1], "Phenomenal consciousness is experience; the phenomenally conscious aspect of a state is what it is like to be in that state. The mark of access-consciousness, by contrast, is availability for use in reasoning and rationally guiding speech and action." Block feels that it is *possible* to have phenomenal consciousness and access consciousness independently of each other, but in general they do interact.

There is no generally agreed upon way of categorizing different types of consciousness. Block's distinction between phenomenal consciousness and access consciousness tries to distinguish between conscious states that either do or do not directly involve the control of thought and action.

Phenomenal consciousness. According to Block, phenomenal consciousness results from sensory experiences such as hearing, smelling, tasting, and having pains. Block groups together as phenomenal consciousness **experiences** such as sensations, feelings, perceptions, thoughts, wants and emotions. Block excludes from phenomenal consciousness anything having to do with cognition, intentionality, or with "properties definable in a computer program".

Access consciousness. Access consciousness is available for use in reasoning and for direct conscious control of action and speech. For Block, the "report ability" of access consciousness is of great practical importance. Also, access consciousness **must** be "representational" because only representational content can figure in reasoning. Examples of access consciousness are thoughts, beliefs, and desires.

A potential source of confusion is that some phenomenal consciousness is also representational. The key distinction to keep in mind about representational content that Block would place in the access consciousness category is that the reason it is placed in the access consciousness category is because of its representational aspect. Elements of phenomenal consciousness are assigned to the phenomenal consciousness category because of their phenomenal content.

Reaction

An immediate point of controversy for Block's attempt to divide consciousness into the subdivisions of phenomenal consciousness and access consciousness is that some people view the mind as resulting (in its entirety) from fundamentally computational processes. This computational view of mind implies that ALL of consciousness is "definable in a computer program", so Block's attempt to describe some consciousness as phenomenal consciousness cannot succeed in identifying a distinct category of conscious states. This viewpoint is highly contentious however, see The problem of machine and digital consciousness for a discussion

As mentioned above, Block feels that phenomenal consciousness and access consciousness normally interact, but it is possible to have access consciousness without phenomenal

consciousness. In particular, believes that zombies are possible and a robot could exist that is "computationally identical to a person" while having no phenomenal consciousness. Similarly, Block feels that you can have an animal with phenomenal consciousness but no access consciousness.

Block shares Chalmers' belief that we can have conscious experiences that are not possible to produce by any type of computational algorithm and that the source of such experiences is "the hard problem" of consciousness. To functionalists Block's position with respect to consciousness is analogous to that of Vitalists who defined Life as being in a category distinct from all possible physical processes. To those who support phenomenal consciousness the functionalist viewpoint is like believing in a flat earth, flat earthers see the world through biblical cosmology and functionalists view it through nineteenth century science. Biologists refute Vitalism by describing the physical processes that account for Life. Cosmologists refute biblical cosmology by describing modern physics. In order to refute Block's claim about the distinction between phenomenal consciousness and access consciousness, it is up to biologists and artificial consciousness researchers to describe computational algorithms that account for consciousness. In order to refute functionalism philosophers and scientists draw attention to the fact that they are trying to explain an internal state of a conscious observer, something that cannot be explained in terms of the external behaviour of machines.

Why are some neurobiologists and computer scientists sure that Block's division of consciousness is wrong? What is the source of Block's certainty that there are non-computational forms of consciousness? One example of phenomenal consciousness discussed by Block is a loud noise that you do not consciously notice because you are paying attention to something else. Block is sure that you were aware of the noise (phenomenal consciousness) but just not "consciously aware" (access consciousness). Many scientists would say that in this case, you were not "consciously aware" of the noise, but it is almost certain that portions of your unconscious brain activity responded to the noise (you could electrically record activity in the primary auditory cortex that is clearly a response to action potentials arriving from the ears due to sound waves from the noise). This suggests that Block's controversial "non-computational" category of phenomenal consciousness includes brain activity that others would categorize as being unconscious, not conscious. Some unconscious brain activity can begin to contribute to consciousness when the focus of one's conscious awareness shifts. This suggests that some of what Block calls phenomenal consciousness is brain activity that can either take place outside of consciousness or as part of consciousness, depending on other things that might be going on in the brain at the same time. If so, we can ask why the consciously experienced version of this kind of brain activity is computational while the unconscious version is not. On the other hand many authors (Eddington, Broad, Penrose, McFadden, Zeh etc) would point out that brain activity could be both computational and phenomenal.

Block stresses that he makes use of introspection to distinguish between phenomenal consciousness and access consciousness. Presumably this means that when the loud noise was not noticed, it was not accessed by introspection. Block has thus defined a category of consciousness that is outside of our "conscious awareness" (although he says we are "aware" of it in some other way) and not accessed by introspection. Maybe it is this inaccessibility of some cases of phenomenal consciousness that motivate Block's idea that such forms of consciousness cannot be computational. When experiences are accessible to

introspection and available for inclusion in reasoning processes, we can begin to imagine computational algorithms for the generation of the content of those experience. However, it is difficult to imagine how the content could become the same as the form of our experience.

Forms of phenomenal consciousness that are open to introspection

In his 1995 article, Block went on to discuss the more interesting cases such as if upon starting to "pay attention to" the load noise (see above) that was previously ignored, the experiencer noticed that there had been *some* earlier experience of the noise, just not of the type that we "pay attention to"; a type of experience that had been just "on the edge" of access consciousness.

In Ned Block's entry for "Consciousness" in the 2004 Oxford Companion to the Mind[2], he discusses another example that he feels distinguishes between phenomenal consciousness and access consciousness.

"Liss[3] presented subjects with 4 letters in two circumstances,

long, e.g. 40 msec, followed by a 'mask' known to make stimuli hard to identify

or

short, e.g. 9 msec, without a mask.

Subjects could identify 3 of the 4 letters on average in the short case but said they were weak and fuzzy. In the long case, they could identify only one letter, but said they could see them all and that the letters were sharper, brighter and higher in contrast. This experiment suggests a double dissociation: the short stimuli were phenomenally poor but perceptually and conceptually OK, whereas the long stimuli were phenomenally sharp but perceptually or conceptually poor, as reflected in the low reportability."

This experiment demonstrates a distinction between

i) reportability of names of the letters

and

ii) perceptual sharpness of the image.

Block's definitions of these two types of consciousness leads us to the conclusion that a non-computational process can present us with phenomenal consciousness of the forms of the letters, while we can imagine an additional computational algorithm for extracting the names of the letters from their form (this is why computer programs can perform character recognition). The ability of a computer to perform character recognition does not imply that it has phenomenal consciousness or that it need share our ability to be consciously aware of the forms of letters that it can algorithmically match to their names.

Reactions

If Block's distinction between phenomenal consciousness and access consciousness is

correct, then it has important implications for attempts by neuroscientists to identify the neural correlates of consciousness and for attempts by computer scientists to produce artificial consciousness in man-made devices such as robots. In particular, Block seems to suggest that non-computational mechanisms for producing the subjective experiences of phenomenal consciousness must be found in order to account for the richness of human consciousness or for there to be a way to rationally endow man-made machines with a similarly rich scope of personal experiences of "what it is like to be in conscious states". Other philosophers of consciousness such as John Searle have similarly suggested that there is something fundamental about subjective experience that cannot be captured by conventional computer programs. This has led to proposals by physicists such as Penrose, Stapp, McFadden etc. for non-digital versions of machines with artificial consciousness.

Many advocates of the idea that there is a fundamentally computational basis of mind feel that the phenomenal aspects of consciousness do not lie outside of the bounds of what can be accomplished by computation[4]. Some of the conflict over the importance of the distinction between phenomenal consciousness and access consciousness centers on just what is meant by terms such as "computation", "program" and "algorithm". In practical terms, how can we know if it is within the power of "computation", "program" or "algorithm" to produce human-like consciousness? There is a problem of verification; can we ever really know if we have a correct biological account of the mechanistic basis of conscious experience and how can we ever know if a robot has phenomenal consciousness? Although of course, such misgivings apply both to those who believe that digital consciousness is possible and those who disagree.

Block's justification of access and phenomenal consciousness uses a nineteenth century idea of the world so cannot be easily sustained against attack from functionalists and eliminativists. However he has clearly described a persistent division in the science and philosophy of consciousness that dates from the time of Aristotle. Aristotle considers this division in terms of those who consider that the soul originates movement and those who consider it to be cognitive, Descartes has the res cogitans and res extensa, Kant has the noumenal and phenomenal, Whitehead has the apparent and causative etc. and even Dennett has the reflex and emergent.

References

1. ^ Block, N. (1995). ON A CONFUSION ABOUT A FUNCTION OF CONSCIOUSNESS. **Behavioral and Brain Sciences** 18 (2): 227-287.

2. ^ Block, N. (2004). "Consciousness" (in R. Gregory (ed.) **Oxford Companion to the Mind**, second edition 2004).

3. ^ Liss, P., (1968). Does backward masking by visual noise stop stimulus processing? **Perception & Psychophysics** 4, 328-330.

4. ^ For a short account, see the Wikipedia entry for phenomenal and access consciousness. Charles Siewert provides a more detailed analysis in his article "Consciousness and Intentionality" in The Stanford Encyclopedia of the Philosophy of Mind.

5. ^ "What is it like to be a bat?" by Thomas Nagel in **The Philosophical Review** LXXXIII, 4 (1974): 435-50.

6. ^ ***On Certainty*** by Ludwig Wittgenstein. Publisher: Harper Perennial (1972) ISBN: 0061316865.

7. ^ Güven Güzeldere described such intuition about the distinctions between phenomenal consciousness and access consciousness as *segregationist intuition*. See "The many faces of consciousness: a field guide" in **THE NATURE OF CONCIOUSNESS; PHILOSOPHICAL DEBATES** Publisher: The MIT Press (1997) ISBN: 0262522101.

Francis Crick (1916 - 2004)

Francis Crick (1994) The Astonishing Hypothesis. The Scientific Search for the Soul. Simon & Schuster Ltd. London.

Crick begins this book with a statement about his opinion of the insignificance of human beings:

"The Astonishing Hypothesis is that "You", your joys and your sorrows, your memories and your ambitions, your sense of personal identity and free will, are in fact no more than the behaviour of a vast assembly of nerve cells and their associated molecules. As Lewis Carroll's Alice might have phrased it: "you're nothing but a pack of neurons". This hypothesis is so alien to the ideas of most people alive today that it can truly be called astonishing."

Crick is not a philosopher so might be forgiven the derogatory "no more than..", as a scientist he realises that the assembly of nerve cells that form a brain is highly complex and difficult to understand.

He suggests that the hypothesis is "so surprising" for three reasons:

"The first is that many people are reluctant to accept what is often called the "reductionist approach" - that a complex system can be explained by the behaviour of its parts and their interactions with each other."

.....

"The second reason why the Astonishing Hypothesis seems so strange is the nature of consciousness. We have, for example, a vivid internal picture of the external world. It might seem a category mistake to believe this is merely another way of talking about the behavior of neurons, but we have just seen that arguments of this type are not always to be trusted."

.....

"The third reason why the Astonishing Hypothesis seems strange springs from our undeniable feeling that Free Will is free. ... I believe that if we solve the problem of awareness (or consciousness), the explanation of Free Will is likely to be easier to solve."

Crick believes that many phenomena in the brain are "emergent" with the vague implication that consciousness may also be emergent. He defines this term in the following way:

"The scientific meaning of emergent, or at least the one I use, assumes that, while the whole may not be the simple sum of the separate parts, its behavior can, at least in principle, be understood from the nature and behavior of its parts plus the knowledge of how all these parts interact."

He wants to avoid the philosophical debates about the nature of consciousness:

"1. Everyone has a rough idea of what is meant by consciousness. It is better to avoid a precise definition of consciousness because of the dangers of premature definition.

"Footnote: If this seems like cheating, try defining for me the word gene. So much is now known about genes that any simple definition is likely to be inadequate. How much more difficult, then, to define a biological term when rather little is known about it."

This is an odd standpoint because any brief review of the ideas of philosophers shows that a good deal is known about phenomenal consciousness. The problem lies in explaining such a bizarre experience, not in defining it.

He then elaborates a further four points covering general features of consciousness and avoiding various types of speculation about consciousness. Excluded are: "what consciousness is for", speculations about consciousness in lower animals and the "self-referential aspect of consciousness"; included are the concept of consciousness in "higher mammals".

As a guide for the scientific investigation of consciousness he puts forward three basic ideas:

"1. Not all the operations of the brain correspond to consciousness.

2. Consciousness involves some form of memory, probably a very short term one.

3. Consciousness is closely associated with attention."

The operations of the brain that do correspond to consciousness are the "neural correlates of consciousness" a term that probably predates Crick's work. Crick shows the openness of ideal science when he concludes with:

"The Astonishing Hypothesis may be proved correct. Alternatively some view closer to the religious one may become more plausible. There is always a third possibility: that the facts support a new, alternative way of looking at the mind-brain problem that is significantly different from the rather crude materialistic view many neuroscientists hold today and also from the religious point of view."

David J Chalmers

Review of "The Conscious Mind: In Search of a Fundamental Theory". Oxford University Press. 1996.

Chalmers is perhaps most famous for his "hard problem" of consciousness:

"... I find myself absorbed in an orange sensation, and *something is going on*. There is something that needs explaining, even after we have explained the process of discrimination: there is the *experience*."p xii

...."This might be seen as a Great Divide in the study of consciousness. If you hold that an answer to the "easy" problems explains everything that needs to be explained, then you get one sort of theory; if you hold that there is a further "hard" problem then you get another."p xiii

Chalmers describes mind as having "phenomenal" and "psychological" aspects.

"At the root of all this lie two quite distinct concepts of mind. The first is the *phenomenal* concept of mind. This is the concept of mind as conscious experience, and of a mental state as a consciously experienced mental

state. ... The second is the *psychological* concept of mind. This is the concept of mind as the causal or explanatory basis for behaviour." p11

Chalmers proposes that consciousness can be explained by a form of "Naturalistic Dualism" that is supported by the following argument:

"In particular, the failure of logical supervenience directly implies that materialism is false: there are features of the world over and above the physical features. The basic argument for this goes as follows: 1. In our world there are conscious experiences. 2. There is a logically possible world physically identical to ours, in which the positive facts about consciousness in our world do not hold. 3. Therefore facts about consciousness are further facts about our world, over and above the physical facts. 4. So materialism is false.

Chalmers describes his naturalistic dualism:

"The dualism implied here is instead a kind of *property* dualism: conscious experience involves properties of an individual that are not entailed by the physical properties of that individual. Consciousness is a *feature* of the world over and above the physical features of the world. This is not to say that it is a separate "substance"; the issue of what it would take to constitute a dualism of substances seems quite unclear to me. All we know is that there are properties of individuals in this world - the phenomenal properties - that are ontologically independent of physical properties." p125

To substantiate his argument he proposes that "zombie" worlds, in which people would behave like us but not be conscious, are logically possible and that worlds that are physically identical to ours, but where conscious experiences are inverted, are logically possible.

Chalmers' argument about the possibility of zombies runs as follows:

A zombie is defined as "...someone or something physically identical to me (or to any other conscious being), but lacking conscious experiences altogether". Chalmers considers that silicon based devices or an entity based on the population of china could lack conscious experience although being able to perform the same functions as a person. He then makes a logical leap to suggest that these examples show that something physically identical to a conscious person could not be conscious:

"But given that it is conceptually coherent that the group-mind set-up or my silicon isomorph could lack conscious experience, it follows that my zombie twin is an equally coherent possibility."p97

In the inverted spectrum argument Chalmers argues that it is logically possible to imagine a world that is physically identical to ours yet where conscious beings experience an inverted spectrum. This assertion is defended on the basis of the elementary science of colour vision.

Unfortunately, without any definite proposal for how conscious experience is realised it seems premature to declare that the zombie and inverted spectrum arguments are correct. Chalmers approaches the problem of the realization of conscious experience when discussing "information".

Chalmers is aware that phenomenal consciousness includes information that is related to information in the physical world:

"A conscious experience is a realization of an information state; a phenomenal judgement is explained by another realization of the same information state. And in a sense, postulating a phenomenal aspect of information is all we need to do to make sure those judgements are truly correct; there really *is* a qualitative aspect to this information, showing up directly in phenomenology

and not just a system of judgements."p 292

Unfortunately he does not explain what a phenomenal "realization of an information state" means. This leads him to consider any information state as potentially capable of conscious experience. He notes that "We find information everywhere, not just in systems that we standardly take to be conscious." and asks whether a thermostat could be conscious. He poses the question "As we move along the scale from fish and slugs through simple neural networks all the way to thermostats, where should consciousness wink out?".

He answers the objection that there may not be any *room* for consciousness in a thermostat by saying that "If consciousness is not logically supervenient, we should not expect to have to find "room" for consciousness in a system's organization; consciousness is quite distinct from the processing properties of the system". He concludes the thermostat article by declaring that:

"While it *could* be the case that experience winks in at a particular point, any specific point seems arbitrary, so a theory that avoids having to make this decision gains a certain simplicity."

This set of ideas leads to the possibility of panpsychism:

"If there is experience associated with thermostats, there is probably experience *everywhere*: wherever there is a causal interaction, there is information, and wherever there is information there is experience." p297

However, Chalmers states that:

"Personally, I am much more confident of naturalistic dualism than I am of panpsychism. The latter issue seems to be very much open. But I hope to have said enough to show that we ought to take the possibility of some sort of panpsychism seriously..." p299

He then postulates that "Phenomenal properties have an intrinsic nature, one that is not exhausted by their location in an information space, and it seems that a purely informational view of the world leaves no room for these intrinsic qualities.". This leads him to suggest that the world is more than just information, that we "*need* some intrinsic nature in the world, to ground information states". This leads him to propose that:

"So the suggestion is that the information spaces required by physics are themselves grounded in phenomenal and protophenomenal properties. Each instantiation of such an information space is in fact a *phenomenal* (or protophenomenal) realization. Every time a feature such as mass and charge is realized, there is an intrinsic property, or *microphenomenal* property for short. We will have a set of basic microphenomenal spaces, one for each fundamental physical property, and it is these spaces that will ground the information spaces that physics requires." p305

So Chalmers takes the proposal of panpsychism, based on the idea that all information spaces might be conscious, to "ground" the information space. Again, any description of how phenomenal consciousness is actually realized in an information space is missing.

Chalmers' explanation of information seems to mystify it, in physics information is arrangements of things, in maths or digital transmission it is usually arrangements of the same thing. For instance 11011 is an arrangement of ones and zeroes along a line - the information has not replaced reality it is simply a way of using reality to represent something else. As Zurek put it: "there is no information without representation". Hence it is difficult to see why microphenomena should be required to instantiate information when the information is already instantiated.

The concept of information as something that can be transmitted from place to place and also as a property of a substance is at the heart of Chalmer's analysis. He states that:

"We have no way to peek inside a dog's brain, for instance, and observe the presence or absence of conscious experience. The status of this problem is controversial, but the mere *prima facie* existence of the problem is sufficient to defeat an epistemological argument, parallel to those above, for the logical supervenience of consciousness. By contrast there is not even a *prima facie* problem of other biologies, or other economies. Those facts are straightforwardly publically accessible, precisely because they are fixed by the physical facts." p74

The patterns of things that comprise "biologies" are, according to this, "physical facts". But from the argument about panpsychism above, physical facts are not grounded, they are information that must be instantiated in some way through "microphenomenal" properties. Chalmers seems to be arguing that nothing logically supervenes on the physical because nothing logically supervenes on mind and physical things are mind.

He introduces the idea of *organizational invariance* as the key feature of a conscious system and declares that a set of beer cans could be conscious:

"I claim that conscious experience arises from fine-grained functional organization. More specifically, I will argue for a *principle of organizational invariance*, holding that given any system that has conscious experiences, then any system that has the same fine-grained functional organization will have qualitatively identical experiences. According to this principle, consciousness is an organizational invariant: a property that remains constant over all functional isomorphs of a given system. Whether the organization is realized in silicon chips, in the population of China, or in beer cans and ping-pong balls does not matter. As long as the functional organisation is right, conscious experience will be determined." p249

Chalmers idea of qualia as a flow of beer cans

Experience of blue Experience of red

Conscious experience is a function. It could occur in a system of beer cans. Is this like phenomenal blue?

Does the difference between the qualia depend entirely on spatial arrangement?

It is intriguing that he considers "functional organisation", or the flow of information in the system, to be sufficient to determine consciousness (ie: invariant arrangements of states in space are unnecessary).

See elementary information theory for a discussion of supervenience in information systems.

Part II: The Problem of Consciousness

Defining the Problem

This section presents the empirical idea of consciousness, what consciousness is like before theories are applied to explain it. It is based on descriptions from the section on Historical Ideas (see below).

The definition and description of consciousness

Before embarking on the analysis of consciousness it is important to have a definition of what it is that we are attempting to explain. The article below considers how empiricists have described consciousness. It shows that consciousness is the space, time and content of our minds (where the content contains intuitions and feelings).

Introduction

Empirical descriptions of consciousness have been available in Western literature for centuries and in Eastern literature for millennia. It is often maintained that no-one can define consciousness but there is a large body of literature that gives a clear empirical description of it. Perhaps the claim that no-one can define consciousness is frustration at the fact that no-one can explain consciousness.

Weiskrantz (1988) asserted that "Each of us will have his or her own idea of what, if anything, is meant by consciousness..." and that insisting upon a precise definition would be a mistake. Koch and Crick (1999) stated that "Consciousness is a vague term with many usages and will, in the fullness of time, be replaced by a vocabulary that more accurately reflects the contribution of different brain processes."

But is consciousness really a "vague term" and should we each have our own idea of what it means? The empirical descriptions of Descartes, Kant and others are summarised below under the headings of space, time, qualia and awareness. These descriptions show that consciousness is not a vague term at all.

Space and Time

Kant (1781) argued that our minds must be capable of representing objects in space and time. Without space, objects could not be differentiated and would have no properties. Without representation in time, the concepts of succession of events and simultaneity would be unknown to us. Descartes (1641, Meditation V, 3) was also clear that imaginings and perceptions are experiences where things are arranged in space and time: "In the first place, I distinctly imagine that quantity which the philosophers commonly call continuous, or the extension in length, breadth, and depth that is in this quantity, or rather in the object to which it is attributed. Further, I can enumerate in it many diverse parts, and attribute to each of these all sorts of sizes, figures, situations, and local motions; and, in time, I can assign to each of these motions all degrees of duration." Descartes was, as was so often the case, well ahead of his time by describing continuity and dimensionality, the factors that define his view of space as an actual vector space accessible to mathematical and physical analysis

(See section on Descartes for a full discussion.)

Gregory (1966) also pointed out that we see things as if they are projected into space around us. The idea of projection was implicit in Kant's and Descartes' descriptions, which are from the viewpoint of an observer looking out at contents of experience, but Gregory is explicit (although he believes that explanations based on the projection are absurd).

Kant and Descartes describe consciousness as something extended in time but it is Clay and James who draw this fully to our attention. James (1890) quotes E.R. Clay who coined the term "specious present" to describe how we exist for more than a durationless instant and then goes on to say: "In short, the practically cognized present is no knife-edge, but a saddle-back, with a certain breadth of its own on which we sit perched, and from which we look in two directions into time. The unit of composition of our perception of time is a duration, with a bow and a stern, as it were--a rearward--and a forward-looking end. It is only [p. 610] as parts of this duration-block that the relation of succession of one end to the other is perceived. We do not first feel one end and then feel the other after it, and from the perception of the succession infer an interval of time between, but we seem to feel the interval of time as a whole, with its two ends embedded in it." Notice how James' observer is at an instant but the mind is stretched over time.

James' mental time is probably not the same as physical time. Hermann Weyl, the Nobel prize-winning physicist, wrote that reality is a "four-dimensional continuum which is neither "time" nor "space." Only the consciousness that passes on in one portion of this world experiences the detached piece which comes to meet it and passes behind it, as history, that is, as a process that is going forward in time and takes place in space" (Weyl 1918). In other words consciousness has a way of containing events in the same order as they occur in the world but seems to use a mental time that is different from physical time.

Qualia

Qualia are types of things that occur in conscious experience. The colour purple is a good example of a quale (Tye, 1997). Hume (1739) pointed out of things in the mind that "There is nothing but the idea of their colour or tangibility, which can render them conceivable by the mind", in other words qualia might be the things in the mind rather than attributes. Qualia appear to be exceptional and inexplicable; Churchland (1988) writes "How on earth can a feeling of pain result from ions passing across a membrane?". Descartes (Meditations VI, 6, 1641) clearly describes qualia.

Awareness

Descartes, Locke, Hume, Reid, Kant and most other empiricist authors in this field describe conscious phenomena as if there is an observer in their mind looking out at qualia or feeling qualia in the space and time around about. Descartes and Kant thought that the mind must also contain a conceptualisation or intuition of the meaning of its space, time and content so that the qualia become grouped into objects, the objects into events and the events into meaning and expectation.

As Kant put it, we have "intuitions" about the relations between things. In modern parlance our conscious experience appears to contain the output from an unconscious processor; although Kant's term, "intuition," is a more scientific approach because it is an observation

without assumptions about causes. If the present is extended in time, or "specious" as Clay put it, then many moments are available through which it is possible to apprehend both a question and its answer: the processor can frame the question and provide the answer. The observation that our minds extend through time means that this processor does not need to be recursive to provide the outputs we experience as intuitions (one moment can contain an intuition about another whilst both are in the mind).

Descartes (Meditations VI, 10, 1641) considered the origin of intuitions: "Further, I cannot doubt but that there is in me a certain passive faculty of perception, that is, of receiving and taking knowledge of the ideas of sensible things; but this would be useless to me, if there did not also exist in me, or in some other thing, another active faculty capable of forming and producing those ideas. But this active faculty cannot be in me [in as far as I am but a thinking thing], seeing that it does not presuppose thought, and also that those ideas are frequently produced in my mind without my contributing to it in any way, and even frequently contrary to my will." Descartes suspected that the ideas were formed unconsciously, probably in the brain.

Types of Consciousness

It is sometimes held that there are many types of consciousness, Anthony (2001) lists: phenomenal consciousness, access consciousness, state consciousness, creature consciousness, introspective consciousness and self-consciousness. Anthony takes the view that these are all 'modulations' of the term consciousness and do not mean that there are in fact different types of consciousness. In other words these 'types of consciousness' are modulations of the intuition of content arranged in space and time that is the singular consciousness described by Kant and Descartes. According to this explanation access consciousness is the time extended form of processes in phenomenal consciousness, self-consciousness is the time extended form of bodily processes and inner speech etc.. As an example, if we say a word then think it soundlessly it is evident that inner speech is whole, time extended words coming from the vague direction of the vocal chords (or both ears), when we move a limb much of the whole movement is present in our experience as a set of displacements at the position of the limb and extended through time.

Observations and Denials

There can be little doubt that most descriptions of conscious phenomenology have described the same things although some have used terms such as 'continuity' for time and 'representation' for space. Our conscious experiences are the experience of being an observer that has qualia distributed in space and time around a point. This experience is imbued with intuitions.

Contrary to the views of Weiskrantz and of Koch and Crick there seems to be no need to await a definition of consciousness. It has been described for centuries. So why did these authors feel a need to suspend any definition?

The answer is that over the years there has been no widely accepted theory of how this empirical consciousness could occur. This led certain philosophers such as Ryle (1949) to question whether the description of consciousness was credible.

In most cases this sceptical analysis begins with an explanatory discussion of consciousness

such as: if information travels from the observation to the observer then the observer contains the information so there must be another observer within to observe this second set of information. In this case the conclusion is that this implies an impossible homunculus or Ryle's "ghost in the machine" so observation and observer's cannot occur in the mind. This argument is wrong. The scientific argument should be: the observed form of conscious experience cannot occur if it relies on information transfer, therefore the hypothesis that information transfer is consciousness is wrong and some other explanation is needed. (This means that although the content of consciousness is derived from the senses via signals in neurones, conscious experience is not these signals flowing into a nexus). In science the observation is paramount and cannot be discarded because it conflicts with theory.

The process of discounting an observation when an explanation fails also applies to other aspects of consciousness studies. As Gregory (1988) put it: " 'If you can't explain it deny it' is one strategy for dealing with embarrassing questions such as 'what is consciousness?' ". If we discount these denials then the empirical observations of Kant and Descartes and the other empiricists are the bedrock of consciousness studies and consciousness can indeed be described as an observation containing the space, time and content of our minds (where the content contains intuitions and feelings).

This simple definition of the experience we call consciousness is internally consistent and can be expressed in mathematical language. Consciousness is a multidimensional manifold with vectors pointing towards the centre (the apparent observation point). The content can be both the input and output of processors that are external to the manifold.

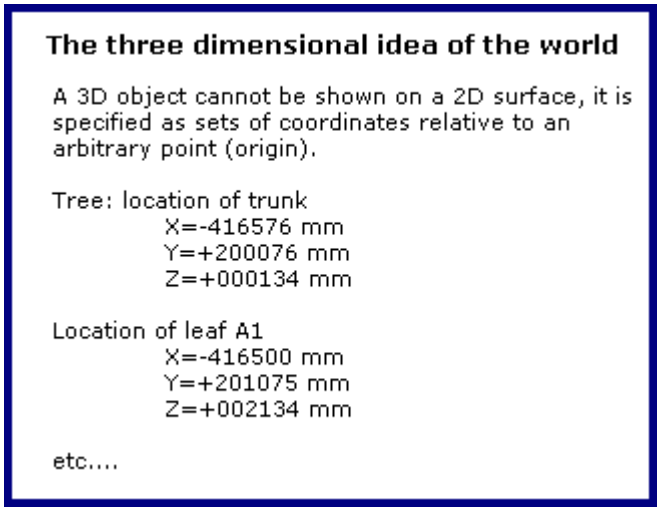
Adapted from the article The description and definition of consciousness by Alex Green in Science and Consciousness Review (with permission of the author).

The viewing point and the observer

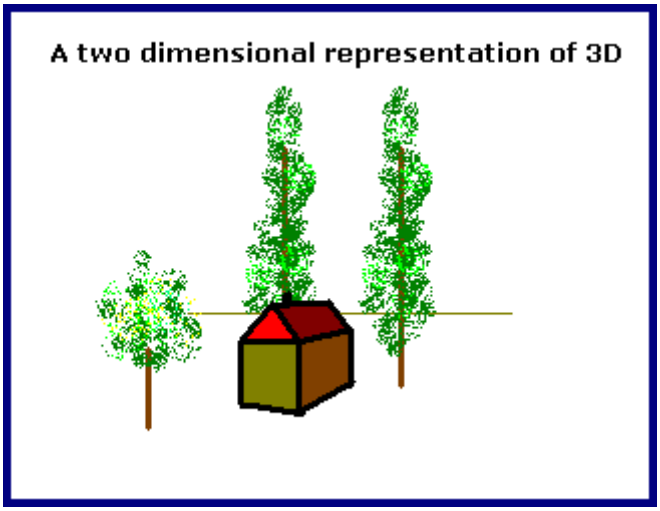
Science begins with empirical descriptions. To experience consciousness simply lean back with your eyes open and listen. Consciousness is the observational space and time that is occurring and the vectors within it that point at the apparent viewing point. It includes bodily sensations, inner speech and the smell on and around things etc. Consciousness is experience itself, it is not usually an experience **of** the content of experience, experience is already there, arranged in space and time (see note below).

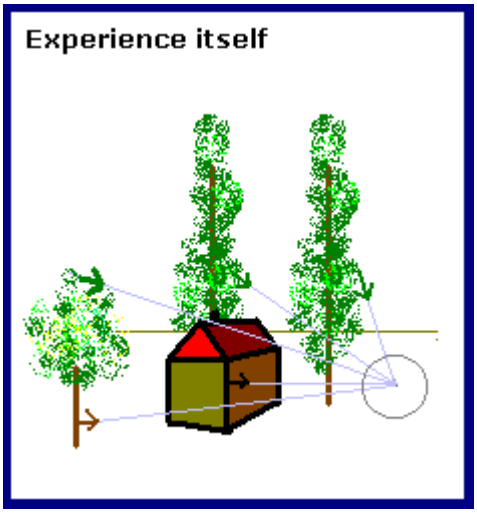
The illustrations show the difference between an actual 3D part of the world, 2D representations of that part of the world, conscious experience itself and naive realism.

It is well known that a 3D object cannot be shown on a 2D surface. Its form is specified as sets of coordinates.



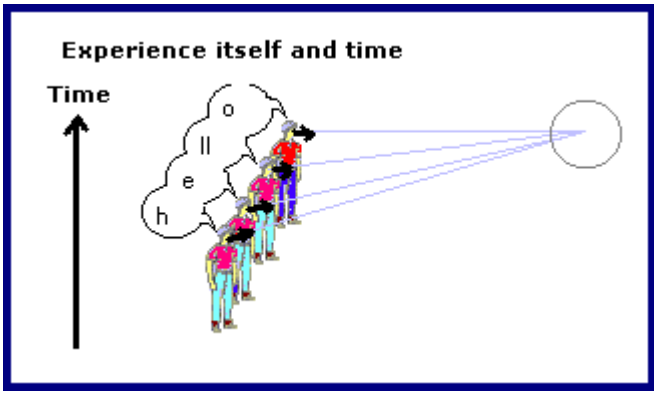
Views are represented on paper using perspective drawings. Pictures that use perspective are scaled images of the world as it would appear on the retina of one eye.





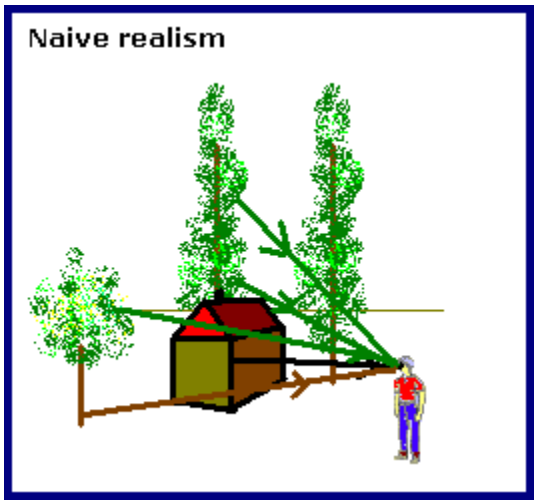
In experience itself things are arranged as vectors directed at a point. Nothing flows into the point. Experience is a manifold of events that are loosely based on data from the retinas and other sense organs. It has contents like the drawing on paper but instead of being a collection of ink particles confined to 2D it is a set of vectors directed at a point.

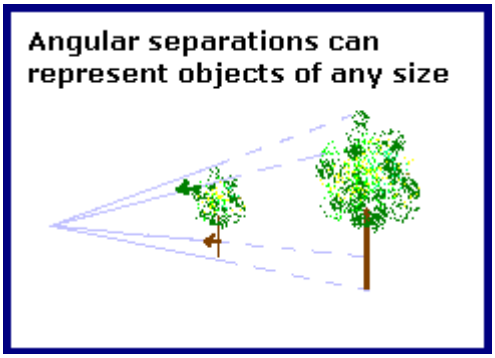
Experience also involves things arranged in time. Things can be simultaneous and there is continuity. Arrangements in time are independent of arrangements in space. The phonemes of a word do not overlap each other and the stages of a movement do not create a smear.



These independent arrangements in time are akin to the way that things that are arranged left and right do not overlap things that are arranged up and down. Left and right are independent of up and down. In a similar way, time seems to be an independent direction for arranging things.

Our experience differs from naive realism. In naive realism experience is believed to be an impossible physical meeting of light rays at a single point in the eye that, through some unspecified mechanism, project back to their source. In experience there seems to be a set of vectors directed at a point. **Naive realism is a primitive dynamical interpretation of experience**, an attempt to explain an empirical geometrical form in terms of flows of matter.





An intriguing feature of the empirical form of experience is that things seem to be separated by angular separations. This allows objects of any size to be represented and explains how a page of text on our laps and the dome of a planetarium can be encompassed in the same form.

The apparent viewing point has caused considerable difficulty for many empiricist philosophers (although the British Empiricists tended to avoid it). When philosophers have stopped describing conscious experience and tried to explain the viewing point they have often resorted to the supernatural: Descartes, Malebranche and Reid all explained the viewing point in terms of a supernatural soul at a point that does the seeing or experiencing. But none of the empiricists describe anything flowing into the viewing point; indeed nothing does flow or could flow into and through a point. The empirical truth is that the viewing point is a geometrical phenomenon, not the recipient of some simultaneous flow of everything in experience. Just look, your viewing point is where everything in experience is directed but things are not pouring into it and it, itself, is a point, it cannot and does not contain anything. This seems to be Aristotle's insight when he wrote "In every case the mind which is actively thinking is the objects which it thinks."

The field of vectors that are the content of consciousness are also difficult to interpret; some philosophers believe that they are in the brain and form a representation of the world whilst others believe that they are directly attached to things in the world beyond the body.

The empirical description of consciousness allows us to make a sharp distinction between the scientific activities of measurement and observation. Measurement is the change in state of a measuring instrument in response to an event in the environment. Observation is the occurrence of events in the geometrical manifold that is our conscious experience.

* Note: the term "experience of" should be reserved for things that act as a source of the content of experience, such as the QM fields that constitute the things that are sensed. We have an "experience of" a flower when signals from the flower are composed into the form of a flower in our experience. Sometimes there is an "experience of" the content of consciousness, for instance when intuitions about the content occur. See later modules for a discussion.

A note on Naive Realism, Materialism and Eliminativism

Naive realism is the belief that the viewing point is a geometrical point where all the light from the things we see is focussed. According to naive realism "seeing" is accomplished by looking from the point back up the light rays to the things in the world. It is a widely held belief even though physics, anatomy and experiment tell us that none of this does or could occur (See the section on vision).

It should be noted that naive realists and Direct Realists may not believe that there is a problem of consciousness because they believe that the form and content of conscious experience is the world itself beyond the body (ie: they believe that the form and content of conscious experience is not a phenomenon based on sensation that happens in the brain). These theories of consciousness (naive and Direct Realism) are discussed in depth in the following modules.

Science begins with observations but there are theories that maintain that our observation should be discarded in favour of theory. This idea comes from a worldview called "nineteenth century materialism" or simply **materialism** which considers that only the present, durationless instant exists and that all things can be explained by flows from place to place. This worldview is taught as school physics. It is an outdated theory that is not the accepted wisdom in modern physics.

Unfortunately materialism is so widely accepted that it leads to highly problematical concepts appearing to be obvious. As an example the extended present of our conscious experience was called "specious" and thought to be obviously wrong by Clay because it is incompatible with materialism (see section on the problem of time). Similarly it is "obvious" that we can only have experiences **of** things rather than containing things because, according to materialism, the present instant has no duration and is immediately non-existent; the extension of things in time can only be conceptualised by endless flows or recursions. The culmination of materialism is Eliminativism and Functionalism in which it is realised that our experience is incompatible with nineteenth century theory and so it is often declared that phenomenal experience is just an illusion that does not exist. As will be seen in the following sections, theories of material flow cannot be used to explain a geometrical manifold.

References

- Anthony,M.V. Is 'Consciousness' Ambiguous? Journal of Consciousness Studies 8(2), 2001, 19-44. http://research.haifa.ac.il/~antony/papers/Ambiguous.htm
- Block, N. 1995. On a Confusion about a Function of Consciousness, The Behavioral and Brain Sciences, 18, 2 (June 1995): 227-247. http://cogprints.ecs.soton.ac.uk/archive/00000231/00/199712004.html
- Churchland, P.S. 1988. Reduction and the neurobiological basis of consciousness. Chapter 14 in Consciousness in Contemporary Science. (Editors: Marcel, A.J. and Bisiach, E.). Oxford Science Publications.
- Descartes, R. 1641. Meditations on First Philosophy http://www.orst.edu/instruct/phl302/texts/descartes/meditations/meditations.html

- James, W. 1890. The Principles of Psychology http://psychclassics.yorku.ca/James/Principles/prin15.htm
- Gregory, R.L. 1966. Eye and Brain. London: Weidenfield and Nicolson.
- Gregory, R.L. 1988. Consciousness in science and philosophy: conscience and con-science. Chapter 12 in Consciousness in Contemporary Science. (Editors: Marcel, A.J. and Bisiach, E.). Oxford Science Publications.
- Kant, I. 1781. The Transcendental Aesthetic. Critique of Pure Reason. Tr. Norman Kemp-Smith. Macmillan Press Ltd. http://www.arts.cuhk.edu.hk/Philosophy/Kant/cpr/cpr-open.html See also this link http://www.utm.edu/research/iep/k/kantmeta.htm
- Koch, C and Crick, F. 1999. 116, Consciousness, neural basis of. http://www.klab.caltech.edu/~koch/Elsevier-NCC.html
- McGinn, C. 1995. Consciousness and Space. In: Conscious Experience, Thomas Metzinger (Ed). 1995, Imprint Academic. http://www.nyu.edu/gsas/dept/philo/courses/consciousness97/papers/ConsciousnessSpace.html
- Ryle, G. 1949. The Concept of Mind. Hutchinson & Co. London.
- Tye, M. 1997. Qualia, Stanford Encyclopedia of Philosophy http://plato.stanford.edu/entries/qualia
- Weiskrantz, L. 1988. Some contributions of neuropsychology of vision and memory to the problem of consciousness. Chapter 8 in Consciousness in Contemporary Science. (Editors: Marcel, A.J. and Bisiach, E.). Oxford Science Publications.
- Weyl, H. 1918. Space Time Matter. Dover Edition 1952, p.217.

The Philosophy of Consciousness

This section is about how regression and recursion seem to undermine the idea of conscious experience being anywhere in the universe..

The conflict - supervenience and the location of the contents of consciousness

When we touch something or look at a view what we are probably touching or seeing is a thing in the world, out there, beyond our bodies. Many philosophers and almost all scientists would agree with this surmise. But is our conscious experience itself directly the things we touch or more like a picture of those things on television or something else entirely?

Conscious experience appears to be a simultaneous set of things (ie: things arranged in space) but where are these things and what is this space? The things that occur in conscious experience could be a virtual reality in the brain based on the world beyond the body, or they could be the things themselves, viewed directly through some unknown phenomenon or it has even been suggested that they could be something non-physical.

This idea of where the contents of conscious experience are located has provoked the fiercest battles in the philosophy of consciousness. There are three broad positions, the first is Direct Realism in which it is held that the contents of conscious experience are directly things in the world, the second is Indirect Realism where it is proposed that the contents of conscious experience are representations, usually in the physical brain, based on things 'out there' in the world and the third is idealism where it is held that there is no physical world, only non-physical conscious experience. These three classifications overlap considerably, for instance some Natural Dualists believe that the contents of sensory experience are directly the world beyond the body but some thoughts are based in a non-physical soul and some philosophers introduce the dualist notion of a "logical space" containing disembodied information.

Philosophers often use the concept of 'supervenience' to examine the location of the contents of consciousness. Supervenience is the relation between two sets of properties. Supervenience can be simple; for example a golden ring supervenes on a piece of the metal gold. Supervenience can also be quite complex such as the idea that life supervenes on the biological processes in a cell. The most difficult cases of supervenience are where a high level description is related to simpler physical properties such as form and content. There are formal statements of supervenience:

```
The properties of A supervene on the properties of B if no two possible
situations are identical with respect to the properties of A while
differing with respect to the properties of B (after Chalmers 1996).
```

Lewis gives a simpler, if less technical, definition of supervenience:

```
A dot-matrix picture has global properties -- it is symmetrical,
it is cluttered, and whatnot -- and yet all there is to the picture
is dots and non-dots at each point of the matrix. The global
properties are nothing but patterns in the dots. They supervene: no
two pictures could differ in their global properties without
differing, somewhere, in whether there is or there isn't a dot".
Lewis, D., 1986, On the Plurality of Worlds, Oxford: Blackwell
```

One set of properties is said to supervene *locally* on another set of properties if the second set is determined by the first. Shape is an example of local supervenience; for instance a gold wire forged in a circle determines a gold ring. A set of properties is said to supervene *globally* on another if the entire context of the properties must be included; for instance, two organisms could be physically identical but demonstrate different behaviours in different environments. In this case the physical form of an organism does not totally determine the behaviour. Philosophers also divide supervenience into logical supervenience and natural supervenience. Logical supervenience deals with possible relations in possible worlds whilst natural supervenience deals with relations that occur in the natural world.

See elementary information theory for a discussion of supervenience in information systems.

A particular problem posed by consciousness studies is whether conscious phenomenal experience supervenes on the physical world and, if so, where. To answer these questions philosophers and neuroscientists must have a good understanding of physics. They should be aware of elementary physical ontology such as kinetic energy being the relativistic mass increase of a particle in a four dimensional universe and Newton's laws being due to the exploration of all paths in space-time. Without a good knowledge of physics there is the danger that we will be asking whether phenomenal consciousness supervenes on an abstract model of the world which does not supervene on the world itself (ie: *we may be asking if conscious phenomenal experience supervenes on Newtonian physics or supervenes on information systems theory rather than asking how phenomenal consciousness might supervene on the natural world*).

(See for instance:

Special relativity for beginners

http://en.wikipedia.org/wiki/Special_relativity_for_beginners

Quantum physics explains Newton's laws of motion

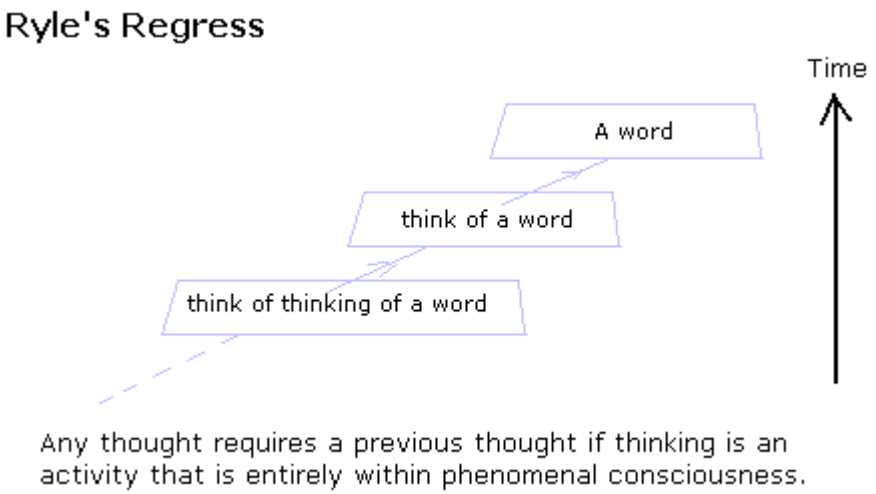
http://www.eftaylor.com/pub/OgbornTaylor.pdf)

The problem of regression

The philosopher Gilbert Ryle was concerned with what he called the intellectualist legend which requires intelligent acts to be the product of the conscious application of mental rules. The intellectualist legend is also known as the "Dogma of the Ghost in the Machine," the "Two-Lives Legend," the "Two-Worlds Story," or the "Double-Life Legend". Ralph Waldo Emerson summarised the intellectualist legend in the statement that "The ancestor of every action is a thought." Ryle argued against the idea that every action requires a conscious thought and showed that this 'intellectualist legend' results in an infinite regress of thought:

> "According to the legend, whenever an agent does anything intelligently, his act is preceded and steered by another internal act of considering a regulative proposition appropriate to his practical problem. [...] Must we then say that for the hero's reflections how to act to be intelligent he must first reflect how best to reflect how to act? The endlessness of this implied regress shows that the application of the criterion of appropriateness does not entail the occurrence of a process of considering this criterion."

(The Concept of Mind (1949))



"The crucial objection to the intellectualist legend is this. The consideration of propositions is itself an operation the execution of which can be more or less intelligent, less or more stupid. But if, for any operation to be intelligently executed, a prior theoretical operation had first to be performed and performed intelligently, it would be a logical impossibility for anyone ever to break into the circle."

Ryle's regress is an adaptation of the regress argument in the philosophy of knowledge (epistemology). In the epistemological regress argument every effect has a cause and every cause must be the effect of a further cause ad infinitum.

Variants of Ryle's regress are commonly aimed at cognitivist theories. For instance, in order to explain the behavior of rats, Edward Tolman suggested that the rats were constructing a "cognitive map" that helped them locate reinforcers, and he used intentional terms (e.g., expectancies, purposes, meanings) to describe their behavior. This led to a famous attack on Tolman's work by Guthrie who pointed out that if one was implying that every action must be preceded by a cognitive 'action' (a 'thought' or 'schema' or 'script' or whatever), then what 'causes' this act? Clearly it must be preceded by another cognitive action, which must in turn must be preceded by another and so on, in an infinite regress unless an external input occurs at some stage.

As a further example, we may take note of the following statement from The Concept of Mind:

"The main object of this chapter is to show that there are many activities which directly display qualities of mind, yet are neither themselves intellectual operations nor yet effects of intellectual operations. Intelligent practice is not a step-child of theory. On the contrary theorizing is one practice amongst others and is itself intelligently or stupidly conducted."

Ryle noted that "theorizing is one practice amongst others." and hence would translate the statement by Emerson into, "The ancestor of every action is an action." or "The ancestor of

every behavior is a behavior". Each behaviour would require yet another behavior to preface it as its ancestor, and an infinite regress would occur.

It should be noted that Ryle's regress is a critique of cognitivism that arises from the Behaviorist tradition. Near the end of The Concept of Mind, Ryle states, "The Behaviorists' methodological program has been of revolutionary importance to the program of psychology. But more, it has been one of the main sources of the philosophical suspicion that the two-worlds story is a myth." But Ryle's brand of logical behaviorism is not to be confused with the radical behaviorism of B. F. Skinner or the methodological behaviorism of John B. Watson. For as Alex Byrne noted, "Ryle was indeed, as he reportedly said, 'only one arm and one leg a behaviorist'."

Arguments that involve regress are well known in philosophy. In fact any reflexive, or self referencing process or argument will involve a regress if there is no external input. This applies whether the agent that engages in the process is a digital computer or intelligent agent (cf: Smith (1986), Yates (1991)).

Ryle's regress suggests that intelligent acts are not created within phenomenal consciousness. They may have non-conscious components or even be entirely non-conscious. Ryle argued that this might mean that consciousness is just a "ghost in the machine" of the brain because consciousness would be epiphenomenal if it is not the creator of intelligent acts. However, as will be seen below, this conclusion may be premature and certainly cannot be used to dismiss phenomenal consciousness as non-existent or not present in the brain.

The Subject-Object paradox

The Subject-Object paradox points out that a conscious subject appears to observe itself as an object. But if it observes itself as an object then, as an object it cannot be a subject.

Wittgenstein gives an example of this paradox:

"5.63 1. The thinking, presenting subject; there is no such thing. If I wrote a book The World as I Found It, I should also have therein to report on my body and say which members obey my will and which do not, etc. This then would be a method of isolating the subject or rather of showing that in an important sense there is no subject: that is to say, of it alone in this book mention could not be made. 5.632. The subject does not belong to the world but it is a limit of the world. 5.633. Where in the world is a metaphysical subject to be noted? You say that this case is altogether like that of the eye and the field of sight. But you do not really see the eye. And from nothing in the field of sight can it be concluded that it is seen from an eye... 5.64 1. ...The philosophical I is not the man, not the human body or the human soul of which psychology treats, but the metaphysical subject, the limit - not a part of the world."(Wittgenstein 1949).

Thomas Reid also uses this paradox to suggest that everything that is observed must be external to the soul. James (1904), Lektorsky (1980) and many others have attempted to resolve the paradox by proposing that there is really no observer, only the observation or 'reflexive act' of perception. These authors have all identified the content of perception with

the world itself to avoid the paradox, however, as will be seen later, there are other solutions to the paradox.

The homunculus argument in philosophy of mind

A Homunculus argument accounts for a phenomenon in terms of the very phenomenon that it is supposed to explain (Richard Gregory (1987)). Homunculus arguments are always fallacious. In the psychology and philosophy of mind 'homunculus arguments' are extremely useful for detecting where theories of mind fail or are incomplete.

Homunculus arguments are common in the theory of vision. Imagine a person watching a movie. He sees the images as something separate from himself, projected on the screen. How is this done? A simple theory might propose that the light from the screen forms an image on the retinas in the eyes and something in the brain looks at these as if they are the screen. The Homunculus Argument shows this is not a full explanation because all that has been done is to place an entire person, or homunculus, behind the eye who gazes at the retinas. A more sophisticated argument might propose that the images on the retinas are transferred to the visual cortex where it is scanned. Again this cannot be a full explanation because all that has been done is to place a little person in the brain behind the cortex. In the theory of vision the Homunculus Argument invalidates theories that do not explain 'projection', the experience that the viewing point is separate from the things that are seen. (Adapted from Gregory (1987), (1990)).

In the case of vision it is sometimes suggested that each homunculus would need a homunculus inside it ad infinitum. This is the **recursion** form of the homunculus concept. Notice that, unlike the case of regress, the recursion would occur after the event.

A homunculus argument should be phrased in such a way that the conclusion is always that if a homunculus is required then the theory is wrong. After all, homunculi do not exist.

Very few people would propose that there actually is a little man in the brain looking at brain activity. However, this proposal has been used as a 'straw man' in theories of mind. Gilbert Ryle (1949) proposed that the human mind is known by its intelligent acts. (see Ryle's Regress). He argued that if there is an inner being inside the brain that could steer its own thoughts then this would lead to an absurd repetitive cycle or 'regress' before a thought could occur:

"According to the legend, whenever an agent does anything intelligently, his act is preceded and steered by another internal act of considering a regulative proposition appropriate to his practical problem."

".... Must we then say that for the ..[agent's].. reflections how to act to be intelligent he must first reflect how best to reflect how to act? The endlessness of this implied regress shows that the application of the appropriateness does not entail the occurrence of a process of considering this criterion."

The homunculus argument and the regress argument are often considered to be the same but this is not the case. The homunculus argument says that if there is a need for a 'little man' to complete a theory then the theory is wrong. The regress argument says that an intelligent agent would need to think before it could have a thought.

If the homunculus argument is applied to the problem of the "intelligent agent" a subtly different result from the regress argument occurs. The homunculus argument applied to Ryle's theory would be phrased in terms of whether the mental attribute of 'reflecting upon things internally' can be explained by the theory that 'the mind is intelligent acts' without the appearance of a homunculus. The answer, provided by Ryle's own logic, is that internal reflection would require a homunculus to prevent it from becoming an infinite regress. Therefore with these assumptions the Homunculus Argument does not support the theory that mind is wholly due to intelligent acts.

The example of Ryle's theory demonstrates another aspect of the Homunculus Argument in which it is possible to attribute to the mind various properties such as 'internal reflection' that are not universally accepted and use these contentiously to declare that a theory of mind is invalid.

The ontological status of regression, recursion and the subject-object paradox

Ryle's regress, when applied to consciousness, is based on an analysis of conscious intellectual activity as a succession of states. At any moment the conscious intellect contains one state such as 'I will think of a word'. This means that either the state has just popped into mind or there was a previous state that gave rise to it such as 'I will think of thinking of a word'. Descartes and other empiricists have noted that thoughts do indeed just pop into mind. So if we transfer Ryle's analysis to the real world we discover that the regress is avoided by removing the starting point of a series of thoughts from conscious phenomenal experience. A train of thought just begins, it has no conscious origin and has probably been synthesised non-consciously.

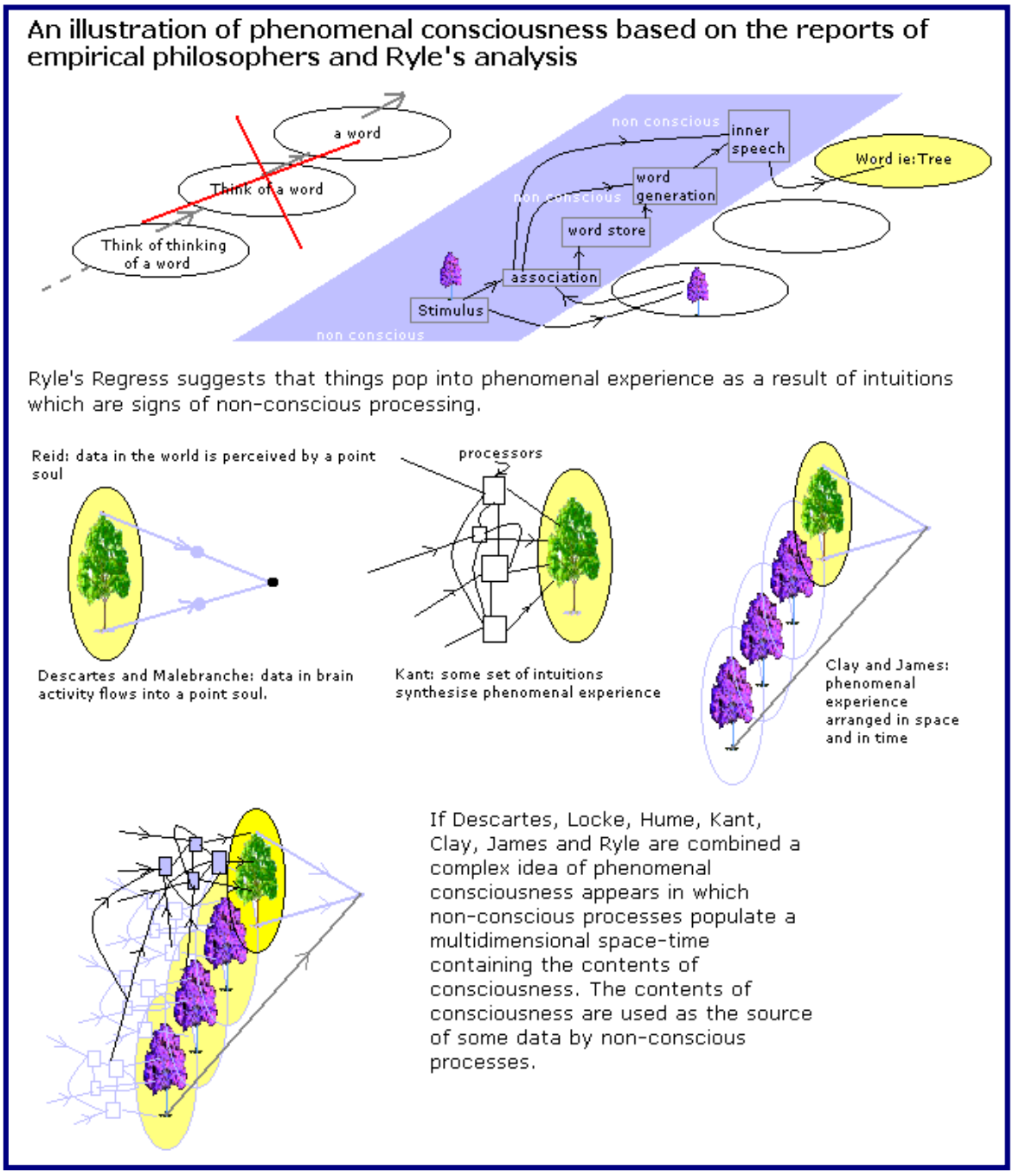
Suppose Descartes and our own experience are correct, suppose thoughts do just pop into mind, if this happens can there still be a conscious intellectual agent or are intellectual agents largely non-conscious? One of the simplest intellectual processes is a test for equality ie: 'does A equal B?' and a routing of flow as a result of the test ie: 'if A = B then goto'. Can an intellectual agent perform an equality test in conscious phenomenal experience?

Consider the test of whether 'A = A', you attend to the left 'A' then the right 'A' and declare them equal. What have you actually done? The feeling that the symbols are equal just pops into mind. Psychologists and philosophers use the word 'intuition' for this popping of answers into mind (Kant 1781). It is usually accompanied by emotional experience (Damasio 1994, Bierman 2004).

If intellectual activity is actually a succession of things that just pop into phenomenal consciousness then Ryle's conclusion that phenomenal conscious is like a "ghost in the machine" of the brain is to some extent justified. Phenomenal consciousness is not intellectual activity. Notice that this was also the conclusion when the Homunculus Argument was applied correctly to intellectual reasoning (see above). Phenomenal consciousness contains the stages, or succession of states, of intellectual activity but does not contain the processes that connect these stages. This observation that conscious experience is a succession of **passive ideas** is well known in philosophy (cf: George Berkeley, Principles of Human Knowledge, 25). That conscious experience should be an orderly arrangement of things, perhaps involved in the stability of brain activity, rather than

the initiator of actions should not surprise anyone: it is observation not action.

The illustration below shows how processing and intuitions are related to conscious phenomenal experience. According to the materialist model (top of picture) conscious experience would be no more than a succession of instantaneous and disconnected ideas. However, the combined empirical descriptions of Aristotle, Descartes, Hume, Kant, Clay and James on phenomenal time and intuition and decision making are also illustrated and it can be seen that the extension in time contained in these descriptions removes the problem.



What Ryle's regress and the recursion argument are telling us is that phenomenal consciousness itself is unlikely to be intelligent acts, processes in a causal chain, although the contents of consciousness could be the succession of states created by such processes. If the contents of consciousness are a succession of states created by non-conscious processes then nothing could flow from place to place within phenomenal consciousness. This would resolve the Subject-Object paradox because the separation between the observation point and the content of consciousness would be due to geometry and time extension, not an impossible dynamical flow of data into the observation point.

References

The problem of regression

- Ryle, G. (1949) The Concept of Mind. The University of Chicago Press, 1949.
- Smith, Q. (1986). The infinite regress of temporal attributions. The Southern Journal of Philosophy (1986) Vol. XXIV, No. 3, 383 (Section 3). http://www.qsmithwmu.com/the_infinite_regress.htm
- Yates, S. (1991). Self referential arguments in philosophy. Reason Papers 16. (Fall 1991) 133-164. http://www.mises.org/reasonpapers/pdf/16/rp_16_7.pdf

The homunculus argument

- Gregory, R.L. (1990) Eye and Brain: The Psychology of Seeing, Oxford University Press Inc. New York.
- Gregory, T.L. (1987). The Oxford Companion to Mind. Oxford University Press.

Subject-object paradox

- James, W. (1904)Does 'Consciousness' Exist? Journal of Philosophy, Psychology, and Scientific Methods, 1, 477-491.
- Hegel, G.W.F. PHILOSOPHY OF MIND: SECTION I. MIND SUBJECTIVE B. PHENOMENOLOGY OF MIND CONSCIOUSNESS Part Three of the Encyclopaedia of the Philosophical Sciences
- Velmans, M. (1996) Consciousness and the "Casual Paradox". Behavioral and Brain Sciences, 19 (3): 538-542.
- [Bermudez, J.L. (1999) The Paradox of Self-Consciousness (representation and Mind) Psycoloquy: 10,#35

Ontological status

- Bierman, D.J. (2004) Non Conscious Processes Preceding Intuitive decisions

http://m0134.fmg.uva.nl/~djb/publications/2004/Intuition_bialsymp2004.pdf

- Damasio, A.R. (1994). Descartes' error: Emotion, reason and the human brain. New York: Grosset/Putnam Book.
- Kant, I. Critique of Pure Reason http://www.arts.cuhk.edu.hk/Philosophy/Kant/cpr/

Direct Realism

- Aydede, M. (2001) Naturalism, introspection, and direct realism about pain. Consciousness and Emotion, Vol. 2, No. 1, 2001, pp. 29-73. http://web.clas.ufl.edu/users/maydede/pain.pdf
- Chapman, C.R., and Y. Nakamura (1999). A Passion of the Soul: An Introduction to Pain for Consciousness Researchers. Consciousness and Cognition, 8: 391-422.
- Dennett, D. (1991). Consciousness Explained. Boston: Little, Brown
- Fowler C A (1986): 'An event approach to the study of speech perception from a direct-realist perspective', J of Phonetics 14(1):3-28.
- Gibson, J. J. (1966) The Senses Considered as Perceptual Systems. Houghton Mifflin Company,Boston.
- Gibson, J. J. (1979) Ecological Approach to Visual Perception.: Lawrence Erlbaum Associates Publishers, Hillsdate.
- Gregory, R.L. 1988. Consciousness in science and philosophy: conscience and con-science. Chapter 12 in Consciousness in Contemporary Science. (Editors: Marcel, A.J. and Bisiach, E.). Oxford Science Publications.
- Le Morvan, Pierre (2004). Arguments against direct realism and how to counter them. *The American Philosophical Quarterly*, *41*(3), 221-234.] (pdf) http://www.tcnj.edu/~lemorvan/DR_web.pdf
- Oliveira, AndrÃ© L. G. and Oliveira, Luis F. (2002) Toward an ecological conception of timbre. In Proceedings Auditory Perception Cognition and Action Meeting 2002, Kansas City. http://cogprints.org/2713/
- Skinner, B. F. Science and Human Behavior . New York: Macmillan, 1953.
- Skinner, B. F. 1971. Beyond Freedom and Dignity. New York: Knopf.
- Skinner, B. F. 1948. Walden Two. New York: Macmillan.
- Tye, M. (2004). ANOTHER LOOK AT REPRESENTATIONALISM ABOUT PAIN. Consciousness and Emotion 2004, special issue on pain, edited by Murat Aydede (with replies by M. Aydede, N. Block, B. Maund, and P. Noordhof http://sun.soci.niu.edu/~phildept/MT/RepresentationalismAndPain.pdf
- Vygotsky, L.S.(1925) Consciousness as a problem in the psychology of behavior. Undiscovered Vygotsky: Etudes on the pre-history of cultural-historical psychology

(European Studies in the History of Science and Ideas. Vol. 8), pp. 251-281. Peter Lang Publishing. http://www.marxists.org/archive/vygotsky/works/1925/consciousness.htm

- Ziff, Paul. "About Behaviourism." Analysis 18 (1958): 132-136. Quoted by Larry Hauser in the Internet Encyclopedia of Philosophy. http://www.utm.edu/research/iep/b/behavior.htm#B. F. Skinner: Radical Behaviorism

Phenomenal consciousness and access consciousness

This part of this section is about the where and when of the experience called consciousness. Is it in the world, in the brain or are the world and brain within it?

Block(1995) drew attention to the way that there appear to be two types of consciousness: *phenomenal* conciousness and *access* consciousness:

> Phenomenal consciousness is experience; the phenomenally conscious aspect of a state is what it is like to be in that state. The mark of access-consciousness, by contrast, is availability for use in reasoning and rationally guiding speech and action. (Block 1995).

See the section on Ned Block's ideas for a deeper coverage of his approach to access and phenomenal consciousness.

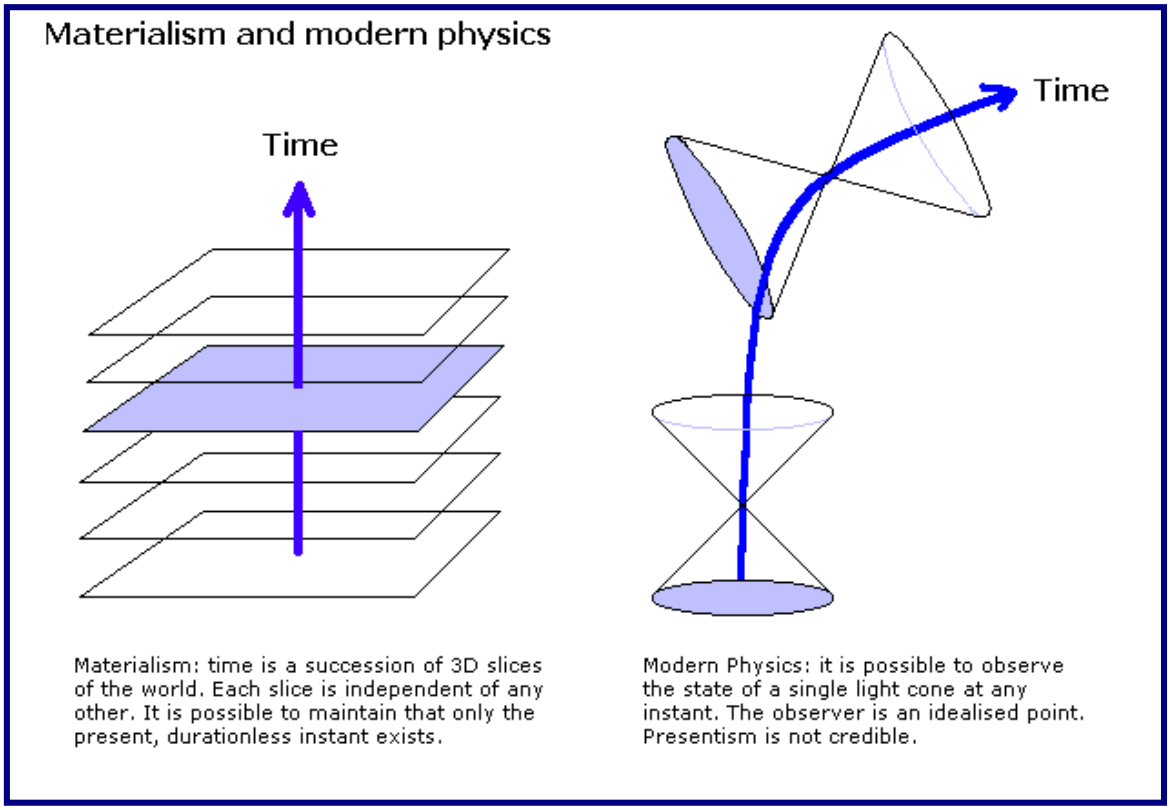
Block uses Nagel's famous (1974) paper, "What is it like to be a bat?" as an exemplary description of phenomenal consciousness. Excellent descriptions have also been proffered by the empiricist philosophers who gave lengthy descriptions of consciousness as partly experience itself. Although Block has formalised the idea of phenomenal and access consciousness similar ideas have also been put forward by many philosophers including Kant and Whitehead.

Access consciousness has two interpretations, in the first, used by Block, it applies to the functions that appear to operate on phenomenal consciousness. In the second, used by the behaviourists and eliminativists, it is some property of the functions of the brain that can be called 'consciousness'.

Curiously there are a significant number of philosophers and neuroscientists who would deny the existence of phenomenal consciousness, who would declare that your experience containing these letters does not exist. The famous twentieth century philosopher Alfred North Whitehead was one of the first to spot that this viewpoint originates in an archaic view of science that is almost religious in its intensity:

> "The eighteenth and nineteenth centuries accepted as their natural philosophy a certain circle of concepts which were as rigid and definite as those of the philosophy of the middle ages, and were accepted with as little critical research. I will call this natural philosophy 'materialism.' Not only were men of science materialists, but also adherents of all schools of philosophy. The idealists only differed from the philosophic materialists on the question of the alignment of nature in reference to mind. But no one had any doubt that the philosophy of nature considered in itself was of the type which I have called materialism. It is the philosophy which I have already examined in my two lectures of this course preceding the present one. It can be summarised as the belief that nature is an aggregate of material and that this material exists in some sense at each successive member of a one-dimensional series of extensionless instants of time. Furthermore the mutual relations of the material entities at each instant formed these entities into a spatial configuration in an unbounded space. It would seem that space---on this theory-would be as instantaneous as the instants, and that some explanation is required of the relations between the successive instantaneous spaces. The materialistic theory is however silent on this point; and the succession of instantaneous spaces is tacitly combined into one persistent space. This theory is a purely intellectual rendering of experience which has had the luck to get itself formulated at the dawn of scientific thought. It has dominated the language and the imagination of science since science flourished in Alexandria, with the result that it is now hardly possible to speak without appearing to assume its immediate obviousness." (Whitehead 1920).

Unfortunately little has changed. The modern generation of computer scientists in particular are taught nothing but nineteenth century materialism and often find it impossible to understand that this is an archaic and discredited idea of science. The difference between materialism and the modern physical view is shown in the illustration below.



As will be seen later, in the discussion on the philosophy of time, there are major differences between the modern and the popular archaic ideas.

According to materialism phenomenal consciousness appears to have no function, it is *epiphenomenal*. It seems to be observation rather than action. This was spotted by Liebniz who proposed that science should be amended because it was in conflict with observation. However, later philosophers such as Huxley in 1874, have taken the view that because phenomenal consciousness appears to have no function according to nineteenth century materialism then it is of no importance or cannot exist.

The idea that phenomenal consciousness cannot exist is a type of *Eliminativism* (also known as Eliminative Materialism). Eliminativism owes much to the work of Sellars(1956) and Feyerbend (1963). Dennett (1978) applied Eliminativism to phenomenal consciousness and denies that pain is real. Others such as Rey(1997) have also applied eliminativism to phenomenal consciousness.

Dennett (1988) redefines consciousness in terms of access consciousness alone, he argues that "Everything real has properties, and since I don't deny the reality of conscious experience, I grant that conscious experience has properties". Having related all consciousness to properties he then declares that these properties are actually judgements of properties. He considers judgements of the properties of consciousness to be identical to the

properties themselves. He writes:

"The infallibilist line on qualia treats them as properties of one's experience one cannot in principle misdiscover, and this is a mysterious doctrine (at least as mysterious as papal infallibility) unless we shift the emphasis a little and treat qualia as logical constructs out of subjects' qualia-judgments: a subject's experience has the quale F if and only if the subject judges his experience to have quale F. "

Having identified "properties" with "judgement of properties" he can then show that the judgements are insubstantial, hence the properties are insubstantial and hence the qualia are insubstantial or even non-existent. Dennett concludes that qualia can be rejected as non-existent:

"So when we look one last time at our original characterization of qualia, as ineffable, intrinsic, private, directly apprehensible properties of experience, we find that there is nothing to fill the bill. In their place are relatively or practically ineffable public properties we can refer to indirectly via reference to our private property-detectors-- private only in the sense of idiosyncratic. And insofar as we wish to cling to our subjective authority about the occurrence within us of states of certain types or with certain properties, we can have some authority--not infallibility or incorrigibility, but something better than sheer guessing--but only if we restrict ourselves to relational, extrinsic properties like the power of certain internal states of ours to provoke acts of apparent re- identification. So contrary to what seems obvious at first blush, there simply are no qualia at all. " (Dennett 1988)

Dennett's reasoning is a classic piece of nineteenth century materialism because it posits that only flows of material are relevant to understanding the world. His assertion that "a subject's experience has the quale F if and only if the subject judges his experience to have quale F" is a statement of the belief that qualia are the same as processes such as judgements. Processes such as judgements are flows of data that do not preserve the form of the content of consciousness. Phenomenal consciousness has both form and content, it is things laid out simultaneously (cf: Plato, Hume), so a process that does not have the form of the thing that it encodes is not the original thing. Judgments are not qualia because they are not congruent with qualia. Dennett's argument is similar to arguing that all cheeses are known by their smell, smells are due to gases therefore cheeses are gaseous.

Dennett's argument has been persuasive and there are now many philosophers and neuroscientists who believe that the problem of phenomenal consciousness does not exist. This means that, to them, what we call 'consciousness' can only be a property of the functions performed by the brain and body. According to these philosophers only access consciousness exists, however, it is a mystery how access consciousness might be experienced if phenomenal consciousness does not exist.

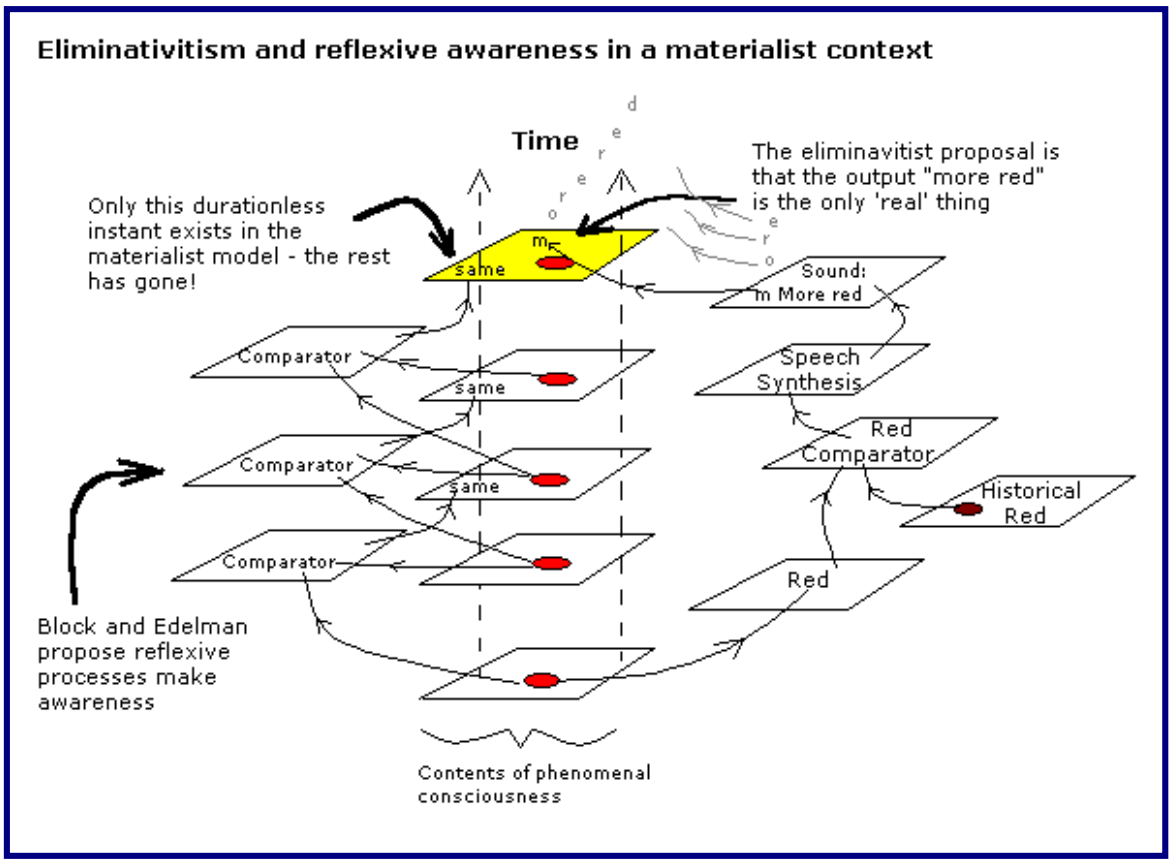
Amongst those who support the idea of phenomenal consciousness there is also a tendency to frame it in terms of nineteenth century theory where one state examines a previous state in a succession over time, for instance Edelman(1993) places the past in memories at an instant and time within experience is explained as continuing modelling processes:

"Primary consciousness is the state of being mentally aware of things in the world--of having mental images in the present. But it is not accompanied by any sense of a person with a past and a future.... In contrast, higher-order consciousness involves the recognition by a thinking subject of his or her own acts or affections. It embodies a model of the personal, and of the past and the future as well as the present. It exhibits direct awareness--the noninferential or immediate awareness of mental episodes without the involvement of sense organs or receptors. It is what we humans have in addition to primary consciousness. We are conscious of being conscious."

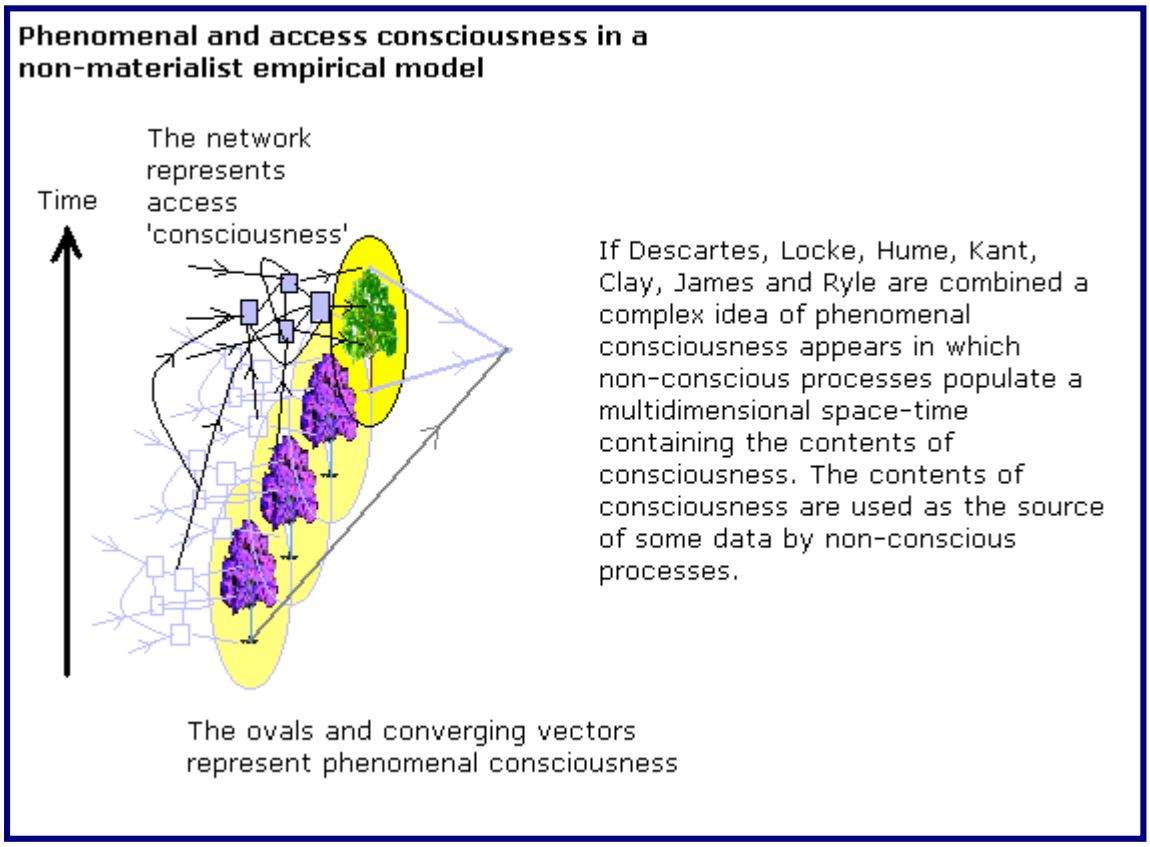
Block(2004) also suggests this flow from state to state in his idea of "Reflexivity" where our idea of familiarity with an object is due to one state being analysed by another:

"Thus in the 'conscious' case, the subject must have a state that is about the subject's own perceptual experience (looking familiar) and thus conscious in what might be termed a 'reflexive' sense. An experience is conscious in this sense just in case it is the object of another of the subject's states; for example, one has a thought to the effect that one has that experience. The reflexive sense of 'consciousness' contrasts with phenomenality, which perhaps attaches to some states which are not the objects of other mental states. Reflexive consciousness might better be called 'awareness' than 'consciousness'. Reflexivity is phenomenality plus something else (reflection) and that opens up the possibility in principle for phenomenality without reflection. For example, it is at least conceptually possible for there to be two people in pain, one of whom is introspecting the pain the other not. (Perhaps infants or animals can have pain but don't introspect it.) The first is reflexively conscious of the pain, but both have phenomenally conscious states, since pain is by its very nature a phenomenally conscious state. "

Both Block and Edelman allow phenomenal consciousness, our experience, as an unexplained phenomenon. Block, Edelman and also Dennett's ideas of consciousness are shown in the illustration below:



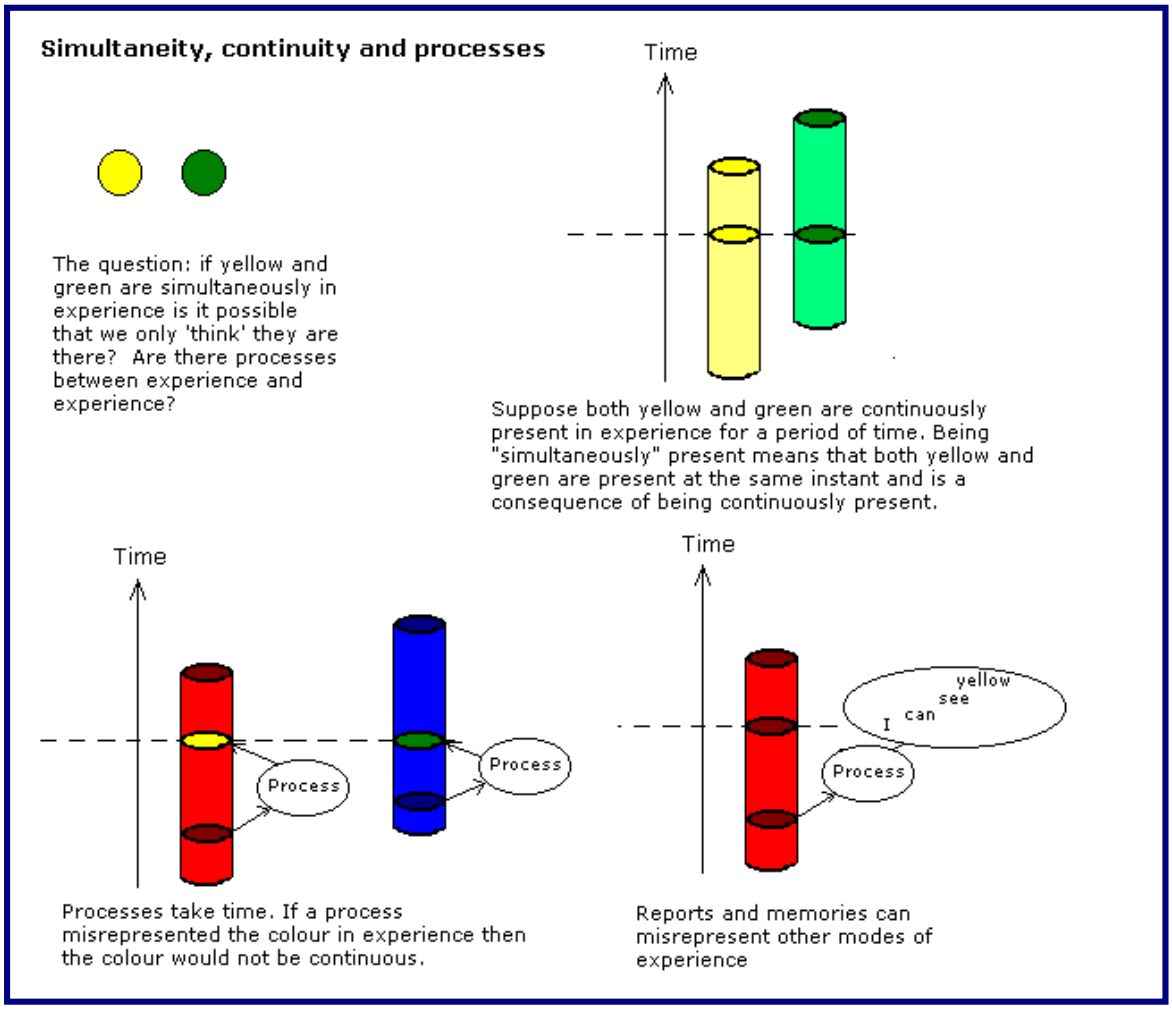
However, if empirical studies are accepted as valid observations, without the artificial constraint of materialism, phenomenal and access consciousness have slightly different meanings, as shown in the illustration below:



According to the empirical reports the present moment in our experience is extended so the succession of outputs or stages of access consciousness could constitute the contents of phenomenal consciousness. In other words phenomenal consciousness is composed of periods of access consciousness. This is how it seems to the empiricist and in our own experience but how such a state could be explained in terms of brain activity is highly problematical. Given that nineteenth century ideas cannot explain such a state a scientific explanation will be required.

The idea that phenomenal consciousness misrepresents or "misdiscovers" itself (Dennett 1988) deserves further discussion. According to materialism the present instant has no duration so can only be known in succeeding instants. Does this mean that experience misrepresents itself? It is interesting that both the materialist and non-materialist paradigms usually require that things which exist have a *duration*. In both paradigms no thing would exist for no time at all. This implies that things have continuity. More than one thing can be continuously present which means that simultaneity and hence space occurs (space being things arranged simultaneously). During the continuity of the thing in experience there cannot be misrepresentation in that mode of experience because, by definition, the thing in experience is the continuous thing. Although things in experience cannot be other things whilst they are continuously present they can be misjudged, misreported and mis-

remembered. This discussion is illustrated below:



Materialism is no longer the conventional wisdom in science, it has been found that time exists and that reality is not a succession of disconnected instants so it is now valid to speak of the extended present of our phenomenal experience as a possible observation rather than as a specious illusion.

Direct Realism

Direct Realism proposes that phenomenal experience is directly objects in the world without any intervening representation. It is motivated by the belief that the Problem of Regression, the Subject-Object Paradox and the recursion form of the Homunculus arguments show that phenomenal consciousness cannot occur in the brain alone. Direct Realists reason that if phenomenal consciousness cannot be things in the brain then it must be something outside the brain.

There are two principle types of Direct Realism: Natural Dualism and Behaviourism (both Radical and Analytical). Thomas Reid is generally regarded as the founder of Direct Realism. In his Natural Dualism he proposed that the soul is in direct contact with the

contents of experience and these contents are things in the world beyond the body. The Direct Realism of Reid is summarised in the statement of his famous disciple Sir William Hamilton: "In the simplest act of perception I am conscious of myself as the perceiving subject and of an external reality as the object perceived". Reid's Natural Dualism has now been largely replaced by radical and analytical behaviourism which eschew the idea of a soul and propose that phenomenal consciousness, if it exists at all, is a behavioural reflex.

The modern justification of Direct Realism mainly consists of arguments against Indirect Realism or Representationalism. Philosophers such as Austin (1962) and Le Morvan (2004) have summarised the Direct Realism debate and have identified the following arguments in favour of Indirect Realism and given rebuttals to each of them:

1. The Causal Argument: perception involves a succession of causal events such as the reflection of photons, bleaching of retinal pigments etc. so perception must involve the end of this causal chain. The Direct Realist response is that, although there may be a causal chain in sensation this does not inevitably imply that the end of the chain is the content of phenomenal experience.

2. The Time Lag Argument: it takes time for light to travel from an object to the senses, time for chemical changes in the retina etc... The Direct Realist response is that direct perception may be referred back in time.

3. The Partial Character of Perception Argument: we only perceive the surface of objects, and then only a part of the surface. As the whole object would be perceived directly perception must be indirect. The Direct Realist response is that direct perception could occur even if only parts of an object were perceived.

4. The Perceptual Relativity Argument: things appear to be different shapes depending upon the point of view. The Direct Realist reponse is that if perception can occur backwards in time it should have no problem occurring back down a line of sight. However, Le Morvan's argument does not seem to encompass the geometrical nature of phenomenal experience, seeking to explain geometry in terms of movement.

5. The Argument from Perceptual Illusion: A stick may appear bent when projecting from the surface of water. Direct Realists apply the argument used in (4) to this problem. The bent stick illusion is a physical event in the world beyond the eye rather than a normal optical illusion such as the Muller-Lyer illusion etc., see (6) for a discussion of optical illusions.

6. The Argument from Hallucination: Hallucinations are not in the world beyond the body. This is highly problematical for Direct Realists expecially when phenomena such as lucid dreams, dreams and visual imaginations are included along with hallucinations. Direct Realists classify these phenomena as not being perceptions or deny that they actually exist as phenomena. Indirect Realists would maintain that all of perception is a reconstruction and use optical illusions such as the Muller-Lyer, Ames Room etc. to justify this contention so the Direct Realist approach to hallucination, dreams etc. might seem like an unwillingness to accept Indirect Realism rather than an argument.

7. The Dubitability Argument (cf: Indubitability argument): we cannot doubt current phenomenal experience but we can doubt the world beyond the body therefore phenomenal experience is not the world beyond the body. Direct Realists fall back on Presentism or

functional Presentism to defeat this argument. If phenomenal experience is instantaneous and made anew at each instant then anything can be doubted.

The points above have summarised the Direct Realist stance on visual perception. Other sensory modalities have also been considered in the Direct Realism debate.

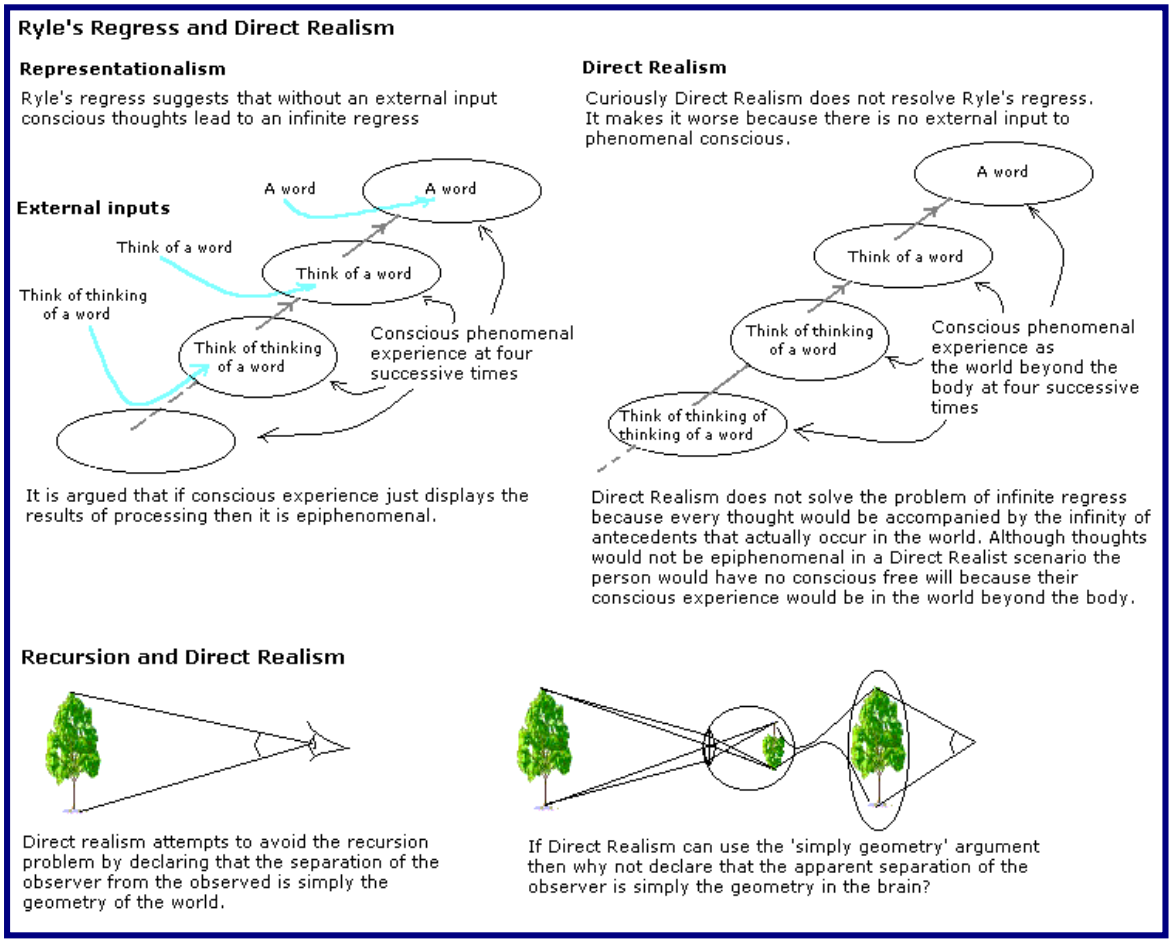
Fowler (1986) considered that sounds were attached to objects in the world. This idea is strange because sounds only seem to be closely attached to objects in the world when these objects are seen as well as heard. For example, when a subject is blindfolded it is found that there can be a large error in locating the position of a sound in the world, this is especially true for low frequency sounds. The Direct Realist approach has difficulty explaining the transition from sounds with an indefinite location when a subject is blindfolded to sounds that are bound to visual events when the blindfold is removed. It also runs into problems explaining how the sound of speech from a single loudspeaker can become bound to lip movements on a cinema screen. If the binding does not occur in the brain then where does it occur?

Pain is particularly problematic for Direct Realism because, unlike colour vision where 'red' is inferred to be a property of electrons or light, pain is an inner experience that is not a property of tissue damage. Tissue damage has properties such as bleeding, wheal formation etc. but pain seems to be phenomenal experience in the brain and 'phantom pain' can occur without tissue damage (see Aydede (2001), Tye (2004) and Chapman, and Nakamura (1999) for further analysis). On closer inspection other sensations also appear to be inner experiences rather than direct sensations. For instance, the red crosses of different hues in the illustration below are all due to the same physical wavelengths of light. In this case the range of hues in experience is unrelated to the actual physical red on the page or screen.

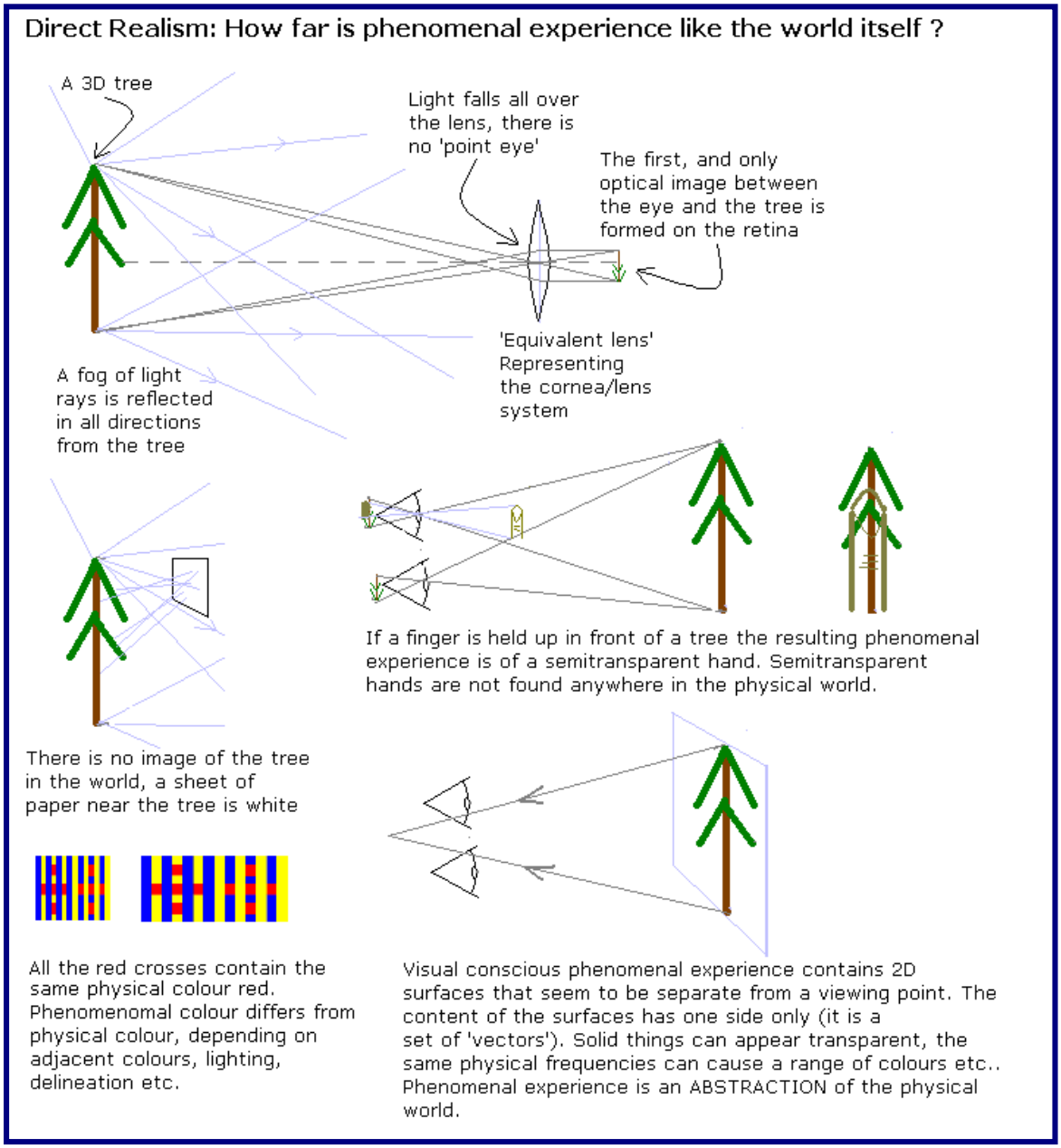
Another problem for Direct Realism is that it does not overcome the problems that it is supposed to solve. The argument for Direct Realism begins with the idea that there are severe problems with representationalism (the idea that phenomenal experience is in the brain) and that direct perception is an alternative that does not have these problems. However on closer inspection Direct Realism suffers from almost same problems as representationalism. If phenomenal experience is the world itself then Ryle's regress applies to the world itself and this can only be avoided by assuming that phenomenal experience is a subset of the world (ie: a representation) that receives input from other parts of the world that are not part of phenomenal experience.

It is also commonly assumed that Direct Realism avoids the recursion argument because it is believed that the separation of the observer from the things that are observed is simply due to the geometry of the world. If such simple geometry is possible between the eye and the world then it should also be possible in the brain and a similar geometrical explanation could be invoked to avoid recursion in representations.

These points are shown in the illustration below:



Scientists have a further problem with Direct Realism. The illustration below demonstrates that our scientific knowledge of the world differs markedly from our phenomenal experience.

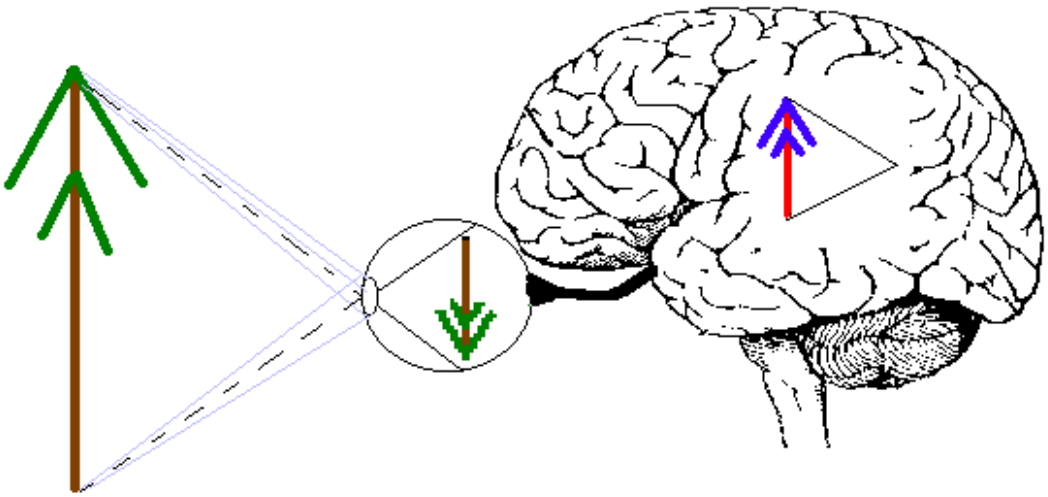


It is difficult to see how the form and content of phenomenal experience could supervene directly on the world beyond the body. The world inferred from measurements beyond the body seems to be a nebulous set of quantum phenomena that are arranged as probability fields in three dimensions at any instant. The objects in this real world are mostly space. The world of phenomenal experience on the other hand contains objects that are one-sided, and are like a 2 dimensional field of vectors directed simultaneously at an observation point which is apparently separate from them. Phenomenal experience is not three dimensional, the rear of objects is not available within it at any instant. Visual phenomenal experience seems to be a geometrical relationship between an abstract observation point and the reflection properties of the part of the world external to the body. It is a form that crudely overlies the angular separations in inferred reality, providing approximate directional data. It is not like things in themselves beyond the body, not even in type, being a set of directional vectors. (See the module on the neuroscience of perception for a discussion of depth perception).

If the form and content of visual phenomenal experience are abstractions separated according to the angular positions of things in the world beyond the body then theories which propose that phenomenal experience is the world itself are problematical. It should be noted that things arranged according to angular positions can appear to overlie any group of similar things along a radius from the centre point.

The View: If it is accepted that experience is a 'view' then any form can overlie any other.

If it is just accepted that the form of the view is due to world geometry then anything along the line of sight could represent the object itself. The important measure would be angular separation.



If it is accepted that angular separations are critical then any replica with the same angular separations could represent the object.

If Direct Realists admit that things are as they appear to be, observed according to angular positions at a 'point eye', then any representation of things on the inside of a sphere of any radius would appear similar. The geometry of the 'point eye' is problematical whether the view contains the world itself or a representation of the world; it cannot be the movements of lumps of matter or energy and the point observation cannot be due to lumps all landing at a point. In other words the 'view' is inconsistent with nineteenth century materialism and will require a scientific explanation.

Radical and Analytical Behaviourism tackle the problem of the difference between the world inferred from measurements beyond the body and the phenomenal world by denying phenomenal consciousness and maintaining that access and reflex consciousness are all that exists or is necessary. Radical Behaviourism is an offshoot of psychological behaviourism and was established as a philosophical adjunct to Marxism by Vygotsky and popularised by Burrhus Frederic Skinner (see Skinner 1953). There is another movement in psychological behaviourism which is similar to Radical Behaviourism called Ecological Psychology (see Gibson 1966, 1979). Analytical Behaviourism is a philosophical movement established by Gilbert Ryle (see Ryle 1949).

The core of Analytical and Radical Behaviourism is the assumption that consciousness exists for a durationless instant so that the Dubitability Argument and the Regression and Recursion Arguments can be applied (Ryle 1949, Skinner 1971 and see the sections on Ryle's Regress and the Subject-Object Paradox above). As a result the Direct Realist is able to insinuate that subjects only think that they have had a particular experience (cf: Dennett 1991a). It is intriguing that Eliminativists also maintain that experience is the world itself for instance, an insight into Dennett's idea of the mind is to be found on pages 407-408 of *Consciousness Explained*:

"It seemed to him, according to the text, as if his mind - his visual field - were filled with intricate details of gold-green buds an wiggling branches, but although this is how it seemed this was an illusion. No such "plenum" ever came into his mind; the plenum remained out in the world where it it didn't have to be represented, but could just be. When we marvel, in those moments of heightened self-consciousness, at the glorious richness of our conscious experience, the richness we marvel at is actually the richness of the world outside, in all its ravishing detail. It does not "enter" our conscious minds, but is simply available"

This is a clear description of Direct Realism (although Dennett does not describe himself as a direct realist).

Radical Behaviourism is sometimes described as the dictum that the only psychological events that are of importance are those that occur outside the head. The absurdity of this has led to jokes:

Q: What does one behaviorist say to another when they meet on the street?

A: You're fine. How am I?

Q: What does one behaviorist say to another after sex?

A: That was great for you. How was it for me?

(Ziff 1958)

However Vygotsky, Skinner and other Radical Behaviourists hold that inner behaviour is

possible so that events within the brain can result in reward or punishment. Vygotsky (1925) describes this approach:

"Consciousness is wholly reduced to the transmitting mechanisms of reflexes operating according to general laws, i.e., no processes other than reactions can be admitted into the organism. The way is also paved for the solution of the problem of self-awareness and self-observation. Inner perception and introspection are possible only thanks to the existence of a proprioceptive field and secondary reflexes, which are connected with it. This is always the echo of a reaction."

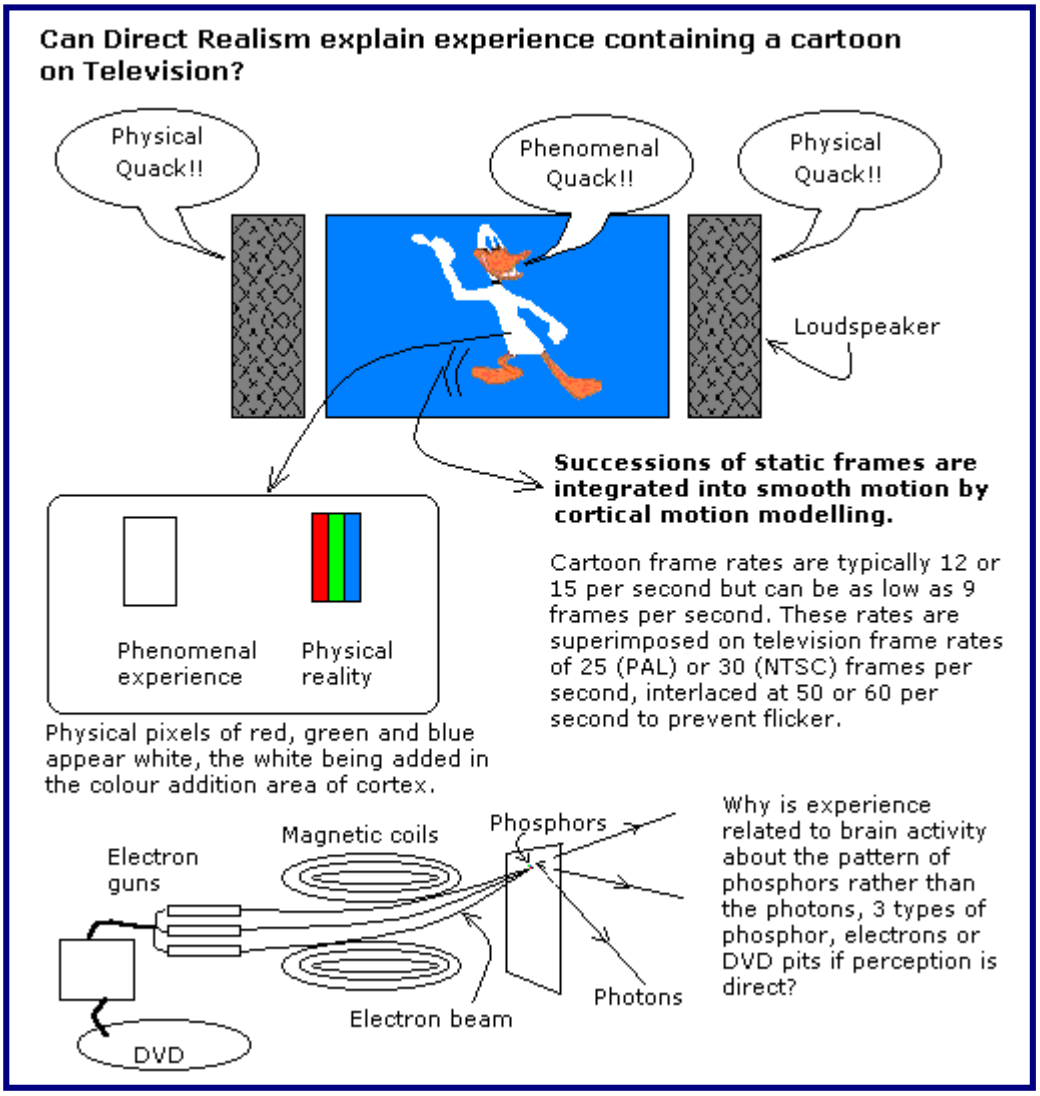
Hence Radical Behaviourists are able to make the claim that what are believed to be representations with phenomenal content are processes. Even events such as pain can then be explained as reflexes involving organs within the skin. However, by opening the possibility that such reflexes could occur at any sense organ, including the eye, this makes Radical Behaviourism a mixed Direct Realist/Indirect Realist philosophy with consciousness as a process, not a separate thing such as phenomenonal consciousness (see the section on representationalism and intentionality below).

But this raises a serious issue for science: can the phenomenal consciousness that seems to contain our observations really be argued out of existence on the basis of a theory? As Gregory (1988) put it: ' 'If you can't explain it â€“ deny it' is one strategy for dealing with embarrassing questions such as 'what is consciousness?' '. But is this the right strategy?

Direct Realism fails to overcome the problems of regression and recursion inherent in representations. It proposes that phenomenal consciousness is identical to the physical world beyond the body but must then use a plethora of arguments to explain why this is evidently not so. When confronted with these problems its proponents resort to the argument that everything can be doubted and can misrepresent itself. Yet it is still widely believed.

It should be noted that Direct Realism is espoused in Religious Natural Dualism, some forms of Augustinian theology, nineteenth century materialism and its offspring such as Marxism, post-modernism, post-Marxism, and various sociological movements. It is also necessary for some forms of Strong AI to occur. Perhaps this explains why few ideas have attracted as much attention and defence as Direct Realism.

It is interesting to compare the Direct Realist and Indirect Realist interpretations of something as simple as a cartoon on television (such as the image below). According to Indirect Realism the cartoon would be a moving representation constructed in the brain using data from the senses. This leads to the prediction of brain mechanisms for modelling motions, combining colours, binding sound and vision etc., many of which have been verified. Can you demonstrate how the theory of Direct Realism could explain the phenomenal experience that contains the cartoon and produce a list of the predictions made by the theory?



In science a theory should be of predictive value, for instance, information theory describes how the state of a thing can be impressed on a carrier so that a signal can be transmitted from one place to another. This theory predicts what will happen when the signal arrives at its destination and how the state of the source can be inferred from the events at the destination, the total amount of information that can be transmitted etc. At the destination it is the form of the signal that is directly known by interaction and measurement, the form of the source is inferred. Direct Realism is a direct challenge to this information theory but does it deliver a more powerful predictive description of phenomenal consciousness or is experience always dependent on what happens to the information flow between things in the world and somewhere in the brain? Does direct realism have a physical theory?

Ultimately it appears as if Direct Realism is about various understandings of Information Theory. For example, Austin (1962) discusses what we see when we see a church camouflaged as a barn and comments that: "We see, of course, a church that now looks like a barn.". Do we see a church or a barn? Scientific information theory is clear about this, the church is an entity composed of selected information from the quantum state of its constituents, the optical image of a camouflaged church is an arrangement of photons emanating from a screen on which it is projected, the retina has an arrangement of chemical

and electrical events based on an optical image and conscious visual experience correlates with the arrangement of things on the retina. The fact that conscious experience also correlates with classifications of the retinal image as a barn or a church suggests that conscious experience is an arrangement of things in the brain based on both the retinal arrangement and the contents of a relational database.

Austin's arguments have been mythologised as a final demonstration that "sense data" theories are false. However, as will be seen below, sense data theories merely claim that there is a succession of information states between an information state outside the body and that reported as conscious experience ie: subjects report that a church is camouflaged when it is camouflaged.

Austin, J.L. (1962) Sense and Sensibilia, ed. by Geoffrey J. Warnock (Oxford, 1962)

Indirect Realism

Indirect realism proposes that phenomenal consciousness exists and is a set of signals or **sense data**, usually in the brain. This was proposed by philosophers from Aristotle to Locke and was probably the most widespread idea of conscious experience until the eighteenth century.

The idea of sense data is discussed in depth by Russell (1912). Russell's original definition is given below:

"Let us give the name of 'sense-data' to the things that are immediately known in sensation: such things as colours, sounds, smells, hardnesses, roughnesses, and so on. We shall give the name 'sensation' to the experience of being immediately aware of these things. Thus, whenever we see a colour, we have a sensation of the colour, but the colour itself is a sense-datum, not a sensation. The colour is that of which we are immediately aware, and the awareness itself is the sensation. It is plain that if we are to know anything about the table, it must be by means of the sense-data -- brown colour, oblong shape, smoothness, etc. -- which we associate with the table; but, for the reasons which have been given, we cannot say that the table is the sense-data, or even that the sense-data are directly properties of the table."

Russell's definition is a materialist concept in which experience is always **of** something because the durationless instant of the present has always gone. As such it differs from some empiricist ideas where experience is not confined to the durationless instant.

Science is Indirect Realist because it holds that the scientist can only make **measurements** of events in the world. These measurements give rise to signals as a result of interaction with the event. According to **decoherence theory** the signals are a state that is a mixture of the state of the measuring instrument and the state of the thing being measured. For example, the eyes are measuring instruments that are sensitive to photons, photons are signals containing a state that is based on the state of electrons in a surface and the state of electrons is based on the state of the surface etc. Scientific inference allows some aspects of the state of the surface to be inferred from the state of the photons.

In modern Indirect Realism there is an attempt to distinguish the phenomenal content of conscious experience from the processing involved in accessing this phenomenal content. According to these theories phenomenal experience is an arrangement of signals that are the content of the experience. This arrangement forms a *representation* of things in the world so this form of indirect realism is known as **Representationalism**. Tye (2003) describes types

of representationalist theory:

"Representationalism, as I have presented it so far, is an identity thesis with respect to qualia: qualia are supposedly one and the same as certain representational contents."

Tye also describes variants of this idea of representationalism:

"Sometimes it is held instead that qualia are one and the same as certain representational properties of experiences; and sometimes it is is argued that these representational properties are themselves irreducible (Siewert 1998). There is also a weaker version of representationalism, according to which it is metaphysically necessary that experiences exactly alike with respect to their representational contents are exactly alike with respect to their qualia. Obviously, this supervenience thesis leaves open the further question as to the essential nature of qualia." (Tye 2003).

In a scientific sense Direct Realists believe that phenomenal experience is the signals that occur next to things in the world beyond the body (which they call "things in themselves") and Indirect Realists usually believe that phenomenal experience is signals in the brain. It can be seen from the pattern of signal flow that the signals travelling into the brain preserve the spatial relationships of the original signals and encode the properties in the original signals. This means that the original signals next to the QM sources and the signals in the brain are equivalent provided the latter are oriented appropriately relative to signals from the body. Either set of signals could transmit or contain the same information. Both Direct and Indirect Realism cannot, at present, explain the physics of how a viewing point occurs in experience ie: how we seem to see through an apparent space to the signals that are the contents of experience. So the choice between Direct Realism and Indirect Realism reduces to whether there is only one set of signals or a chain of signals between the world and phenomenal experience.

The philosophical arguments for Indirect Realism are listed below:

1. Variable perspective: when we see things the view changes so what we see must be a different set of signals depending on the view rather than a constant object.

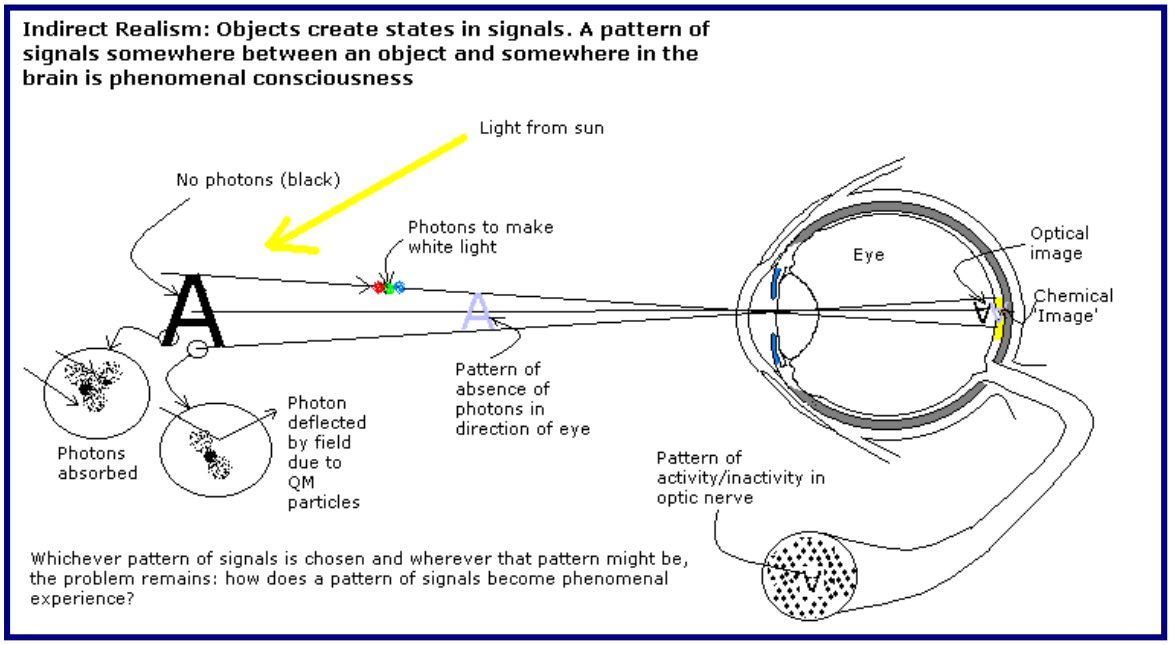
2. Illusions: we can see through fingers and see a variety of colours where measurements tell us one exists. Direct Realists quote the "bent stick illusion", which is not really an illusion at all, being a physical event.

3. Hallucinations: two people can have phenomenal experience containing a table. The first may be viewing a real table whereas the second may be hallucinating a table. If the tables are the same (phenomenally) then experience is indirect.

4. Double vision: press the side of one eye, two images appear (cf Hume 1739) yet there are not two things in the world.

5. Time gap arguments: according to materialism the past has gone. The things being seen no longer exist in the state that relates to the state in experience. In the extreme case, some stars in the night sky no longer exist but are still in experience so experience must be a derived signal.

6. Secondary qualities such as pain, colour and smell do not exist as physical things in the source of signals and are likely to be properties of signals in the brain.



Indirect Realism has received strong support from recent discoveries in neuroscience, for example, it is now clear that both the colour and motion in phenomenal experience are added by cortical processes. In Cerebral **Achromatopsia** patients have suffered trauma to area V4 of the cerebral cortex and report seeing the world in greyscale with no colour vision and in Congenital Achromatopsia people do not even understand the meaning of 'colour'. In an astonishing ailment called **Akinetopsia** patients perceive movement as a succession of stationary images (Rizzo et al 1995). Akinetopsia is usually associated with damage to cortical area V5. Moutoussis and Zeki (1997) have demonstrated that the addition of colour occurs more rapidly than the addition of motion. The section on the Neuroscience of Consciousness describes these discoveries and many other aspects of the creation of phenomenal experience in the brain.

Unfortunately knowledge of the whereabouts of the signals that are the content of conscious experience does not resolve the problem of phenomenal consciousness. Whether these signals are next to objects in the world or at the end of a chain of signals in the brain there still remains the problem of how they become arranged in the form of experience.

Intentionality and representation

There is a materialist interpretation of representationalism in which representations are redefined as intentional states:

"One way of explaining what is meant by 'intentionality' in the (more obscure) philosophical sense is this: it is that aspect of mental states or events that consists in their being of or about things (as pertains to the questions, 'What are you thinking of?' and 'What are you thinking about?'). Intentionality is the aboutness or directedness of mind (or states of mind) to things, objects, states of affairs, events. So if you are thinking about San Francisco, or about the increased cost of living there, or about your meeting someone there at Union Square -- your mind, your thinking, is directed toward San Francisco, or the increased cost of living, or the meeting in Union Square. To think at all

is to think of or about something in this sense. This 'directedness' conception of intentionality plays a prominent role in the influential philosophical writings of Franz Brentano and those whose views developed in response to his (to be discussed further in Section 3)."(Siewert 2003)

This definition allows "representation" to be redefined as a data stream rather than a set of things arranged in some mental or neural state that represents things in space. Husserl thought this approach would allow a description of consciousness that "carefully abstains from affirming the existence of anything in spatio-temporal reality" (Siewert 2003) although it could be argued that a data stream such as any description can never escape the constraints of representation in time at some place.

Unfortunately the concept of "intentionality" has become so diverse that it could be applied to almost any aspect of the description of consciousness. An interesting example of this is given by Loar (2001) where "intentionality" is considered to overlap "representing" and "conceiving":

"A person's thoughts represent things to her -- conceive things -- in many ways: perceptually, memory-wise, descriptively, by naming, by analogy, by intuitive sorting, theoretically, abstractly, implicitly and explicitly. These various manners of conceiving have something in common: they have intentional properties, and they have them essentially.

The usage of the term "intentional state" has become so broad that it now means little more than a state that is about another state.

References

Siewert, C. (2003). Consciousness and Intentionality. The Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/consciousness-intentionality/

Also see Loar, B. (2001) Phenomenal Intentionality as the Basis of Mental Content, in Reflections and Replies, ed. M. Hahn and B. Ramsberg, MIT. http://www.nyu.edu/gsas/dept/philo/courses/concepts/loar.html Williford, K. (2002) The intentionality of consciousness and the consciousness of intentionality. Intentionality: Past and Future, edited by Gabor Forrai and George Kampis, Amsterdam/New York: Rodopi. http://web.stcloudstate.edu/kwwilliford/Intentionality%20of%20Consciousness.pdf

Cartesian materialism

The term "Cartesian materialism" once meant the idea that the mind is in the brain (see for instance Block 1995). The term had largely fallen out of use in philosophy until revived by Daniel Dennett (1991) in the book *Consciousness Explained*. Dennett uses a very particular definition of the term in his discussions and also uses a particular definition of the word "mind". See the section on Daniel Dennett for Dennett's critique. Philosophers who adhere to the idea that the mind is in the brain tend to call themselves "indirect realists" or "representationalists" where the substrate of conscious experience is in the brain and would deny that Dennett's critique applies to their proposals.

Identity theories of mind

The idea that mental states are brain states is known as the **identity theory of mind**. There are two sorts of identity theory, in **type** identity theory it is held that mental states are identical to brain states whereas in **token** identity theory it is held that mental states

correlate with brain states.

Type identity theory was attacked by Putnam in "The Nature of Mental States" where he pointed out that if mental states are functions then type identity theory would presuppose that animals that had the same mental states would need to have identical brain structures. He suggested that this is unlikely, it being more probable that animals have functional systems that perform similar overall functions but which are not identical. In other words, if it is assumed that conscious experience is a set of functions then token identity theory is more probable than type identity theory.

Putnam's critique does not preclude identity theories of mind that involve "passive ideas" (ie: states that are not classical functions).

Most identity theories of mind would be representational, the physical states representing the world in some way. All identity theories of mind involve Cartesian materialism in the sense of the mental states being brain states. According to identity theories the mind is in the brain.

References:

Putnam, H. (1967) The nature of mental states. In The Nature of Mind, edited by Rosenthal, pp. 197-203. Originally published as "Psychological predicates in Art, Mind, and Religion", edited by Capitan and Merill, pp. 37-48.

Dualism

Prior to considering the arguments surrounding dualism it is important to have a clear idea of "information" because many of these arguments have parallels with the difference between information as a set of states that can be transmitted and the substrate on which this information is expressed or from which the information is derived.

Cartesian dualism

Descartes, the founder of Cartesian geometry, analysed his experience and developed an empirical description of how it is arranged. He described mental images and perceptions as extended in space and with a duration. He called these extended things **ideas** (Cartesian ideas) and proposed that they are patterns in the brain. Descartes thought the pineal gland was the most likely location for these ideas because it is one of the few single organs in the brain. He also proposed that there is a rational soul that directly contacts these ideas:

"Now among these figures, it is not those imprinted on the external sense organs, or on the internal surface of the brain, which should be taken to be ideas - but only those which are traced in the spirits on the surface of gland H *(where the seat of the imagination and the 'common sense' is located)*. That is to say, it is only the latter figures which should be taken to be the forms or images which the rational soul united to this machine will consider directly when it imagines some object or perceives it by the senses." Descartes (1664)

See section on Descartes for more information and references.

Descartes considered that the soul was a physical point, an unextended entity that acts like a mind's eye. He called this unextended place the *res cogitans* and concluded that it was a substance that differed from that of material things:

".. I thence concluded that I was a substance whose whole essence or nature consists only in thinking, and which, that it may exist, has need of no place, nor is dependent on any material thing; so that " I," that is to say, the mind by which I am what I am, is wholly distinct from the body, and is even more easily known than the latter, and is such, that although the latter were not, it would still continue to be all that it is."Descartes (1637)

This unextended substance that is not material gives the word "substance" a new meaning. It has been attacked as a concept by Locke, Hume, Berkely and many other philosophers. The concept of there being two substances, that which composes the physical world and that which composes the soul, is the origin of the word **Dualism**.

Cartesian dualism is an attempt to explain our experience. According to Descartes, who only had access to the limited scientific knowledge of the eighteenth century, something supernatural would be needed for an unextended viewing point to exist.

Reid's **Natural Dualism** also has a point soul looking at things but proposes that the things in question are forms in the world rather than in the brain.

Property dualism

Another sort of dualism has arisen out of a particular interpretation of the regress and homunculus arguments. These arguments show that phenomenal experience is not due entirely to flows from place to place (ie: it is not due to classical processes and functions).

As Goldman (1993) pointed out, qualitative experience does not seem to be needed in a functional description of a system:

"For any functional description of a system that is in pain (or has an itch), it seems as if we can imagine another system with the same functional description but lacking the qualitative property of painfulness (or itchiness)."

Certainly a functional system that merely reports the words "I am in pain" when it is dropped on the floor does not require any qualitative property of painfulness. The *absent qualia* arguments suggest that even in a large system there would be no need for qualitative properties for the performance of any classical function.

Chalmers (1993) commenting on Goldman's point, said that this implies that **zombies** might exist, functional replicas of humans but without qualia. He then denied that a complete functional replica of a human could exist without also including qualia:

"It seems to me that the only way to avoid this conclusion is to deny that Zombie Dave is a conceptual possibility; and the only principled way to deny that Zombie Dave is a conceptual possibility is to allow that functional organization is conceptually constitutive of qualitative content." Chalmers (1993).

In other words he identifies qualia with function. According to Chalmers (1996) qualia are a particular type of function:

"I claim that conscious experience arises from fine-grained functional organization. More specifically, I will argue for a principle of organizational invariance, holding that given any system that has conscious experiences, then any system that has the same fine-grained functional organization will have qualitatively identical experiences. According to this principle, consciousness is an

organizational invariant: a property that remains constant over all functional isomorphs of a given system. Whether the organization is realized in silicon chips, in the population of China, or in beer cans and ping-pong balls does not matter. As long as the functional organisation is right, conscious experience will be determined." p249

Chalmers is actually making two major points, firstly that qualia occur during the motion of things (functions), secondly that qualia are independent of any particular substrate.** For the first point to be consistent with materialism the qualia must have no effect on the function, they must be **epiphenomenal**. Epiphenomenal qualia would not be forbidden by the regress and homunculus arguments and would be akin to Berkeley's "passive ideas".

Whether or not epiphenomenal qualia are physical depends upon the definition of the word "physical". If physical functions cause qualia but qualia cannot affect functions then the qualia are "physical" in the sense of being caused by physical events but might be regarded as non-physical in the sense of being isolated from further physical events. In philosophical terms they violate the principle of **Causal Closure**. However, there are other definitions of physicalism based on arguments such as **Methodological Naturalism** which hold that anything that can be investigated using the methods of natural science is a physical thing (see Stoljar 2001). Thus, although epiphenomenal qualia may not conform to materialism they may be encompassed by physicalism; as events that are related to material events they are awaiting a physical theory of how they emerge from a given function.

The reader might consider whether phenomenal consciousness is indeed epiphenomenal. Empirical reports describe it as something that is different from the world beyond the body (see direct realism) - could we generate empirical reports of an epiphenomenon?

The term **property dualism** describes how physical events might give rise to a set of properties that cannot be predicted from the fine structure of the physical system. The "dualism" is present because one set of events is related to two sets of properties, one of which is not related by materialism to the set of events. In the case of the proposal about consciousness outlined above an extra assumption, beyond materialism, would be needed to explain qualia. Property dualism might be defined as a theory that there could be a theory of consciousness but that this requires some new assumption.

As far as the "when and where" of consciousness are concerned, property dualism states that it is somewhere in the processes performed by the organism and the parts of the organism.

** In terms of information processing, Chalmers is proposing that qualia are the enactment of a particular information processing structure.

- Block, N. (2006). Max Black's objection to mind-body identity. Oxford Review of Metaphysics No.3. http://philrsss.anu.edu.au/BlockFest/MaxBlack.rtf
- Chalmers, D.J. (1993) Commentary on "The Psychology of Folk Psychology". Behavioral and Brain Sciences 16:35-36, 1993
- Chalmers, D.J. (1996) "The Conscious Mind: In Search of a Fundamental Theory". Oxford University Press. 1996.
- Goldman, A.I., (1993). The Psychology of Folk Psychology. Behavioral and Brain Sciences 16: 15-28 (1993). http://www.ecs.soton.ac.uk/~harnad/Papers/Py104/goldman.psyc.html

Idealism

According to Idealism only conscious experience truly exists, the physical world is an illusory interpretation of this experience. Descartes pointed out one of the major problems for Idealism, he noticed that even if experience is like a dream it might contain consistent relationships that could be called "science" and hence be indistinguishable from the realist idea of the world.

The form and content of personal conscious experience might be related to the structure of the world and brain in several ways. It could be a **solipsism** or be the mind of God.

*This is a **stub** and needs expansion*.

Panpsychism

According to Panpsychism everything in the world may be conscious; consciousness is a fundamental entity like energy. Conscious experience is then due to an interaction of conscious entities, personal conscious experience being centred on the individual. In some interpretations, such as monadism, Panpsychism and Idealism can overlap because the universe is conceived as being composed of an infinity of point consciousnesses that each contain information about the whole universe.

The form and content of personal conscious experience might be related to the structure of the world and brain in many ways.

*This is a **stub** and needs expansion*.

References

Seager W. & Allen-Hermanson (2001). 'Panpsychism.' Online document: Stanford Encyclopedia of Philosophy, [online], http://plato.stanford.edu/entries/panpsychism/, [accessed 10/2005]

Supervenience

- Chalmers, D.J. (1996). The Conscious Mind. Oxford University Press.

The problem of regression

- Ryle, G. (1949) The Concept of Mind. The University of Chicago Press, 1949.

The homunculus argument

- Gregory, R.L. (1990) Eye and Brain: The Psychology of Seeing, Oxford University Press Inc. New York.
- Gregory, T.L. (1987). The Oxford Companion to Mind. Oxford University Press.

Subject-object paradox

- James, W. (1904)Does 'Consciousness' Exist? Journal of Philosophy, Psychology, and

Scientific Methods, 1, 477-491.

- Hegel, G.W.F. PHILOSOPHY OF MIND: SECTION I. MIND SUBJECTIVE B. PHENOMENOLOGY OF MIND CONSCIOUSNESS Part Three of the Encyclopaedia of the Philosophical Sciences

- Velmans, M. (1996) Consciousness and the "Casual Paradox". Behavioral and Brain Sciences, 19 (3): 538-542.

- [Bermudez, J.L. (1999) The Paradox of Self-Consciousness (representation and Mind) Psycoloquy: 10,#35

Ontological status

- Bierman, D.J. (2004) Non Conscious Processes Preceding Intuitive decisions http://m0134.fmg.uva.nl/~djb/publications/2004/Intuition_bialsymp2004.pdf

- Damasio, A.R. (1994). Descartes' error: Emotion, reason and the human brain. New York: Grosset/Putnam Book.

- Kant, I. Critique of Pure Reason http://www.arts.cuhk.edu.hk/Philosophy/Kant/cpr/

Phenomenal consciousness and access consciousness

- Block, N. (1995) On a confusion about a function of consciousness. Behavioral and Brain Sciences 18 (2): 227-287. http://www.bbsonline.org/documents/a/00/00/04/31/bbs00000431-00/bbs.block.html

- Block, N. (2004). *The Encylopedia of Cognitive Science*. http://www.nyu.edu/gsas/dept/philo/faculty/block/papers/ecs.pdf

- Dennett, D. (1978). Why You Can't Make a Computer that Feels Pain, in: Brainstorms. Cambridge, MA: MIT Press: 190-229.

- Dennett, D. (1988). Quining Qualia. in A. Marcel and E. Bisiach, eds, Consciousness in Modern Science, Oxford University Press 1988. Reprinted in W. Lycan, ed., Mind and Cognition: A Reader, MIT Press, 1990, A. Goldman, ed. Readings in Philosophy and Cognitive Science, MIT Press, 1993. http://ase.tufts.edu/cogstud/papers/quinqual.htm

- Edelman, G.M. (1993). Bright Air, Brilliant Fire: On the Matter of the Mind. New York:

- Feyerabend, P. (1963). Mental Events and the Brain, Journal of Philosophy 40:295-6.

- Huxley, T. H. (1874). On the Hypothesis that Animals are Automata, and its History, The Fortnightly Review, n.s.16:555-580. Reprinted in Method and Results: Essays by Thomas H. Huxley (New York: D. Appleton and Company, 1898).

- Nagel, T. (1974). What is it like to be a bat? The Philosophical Review LXXXIII, 4 (October 1974): 435-50. http://members.aol.com/NeoNoetics/Nagel_Bat.html

- Rey, Georges, (1997). Contemporary Philosophy of Mind.Blackwell: Oxford

- Sellars W. (1956). Empiricism and the Philosophy of Mind, In: Feigl H and Scriven M (eds) The Foundations of Science and the Concepts of of Psychology and Psychoanalysis: Minnesota Studies in the Philosophy of Science, Vol. 1. Minneapolis: University of Minnesota Press: 253-329

- Whitehead, A.N. (1920). The Concept of Nature. Chapter 3: Time. http://spartan.ac.brocku.ca/~lward/Whitehead/Whitehead_1920/White1_03.html

Direct Realism

- Aydede, M. (2001) Naturalism, introspection, and direct realism about pain. Consciousness and Emotion, Vol. 2, No. 1, 2001, pp. 29-73. http://web.clas.ufl.edu/users/maydede/pain.pdf

- Chapman, C.R., and Y. Nakamura (1999). A Passion of the Soul: An Introduction to Pain for Consciousness Researchers. Consciousness and Cognition, 8: 391-422.

- Dennett, D. (1991a). Consciousness Explained. Boston: Little, Brown

- Dennett, D. (1991b). Lovely and suspect qualities. Commentary on David Rosenthal, "The Independence of Consciousness and Sensory Quality" in E. Villanueva, ed., Consciousness, (SOFIA Conference, Buenos Aires), Atascadero, CA: Ridgeview 1991. http://ase.tufts.edu/cogstud/papers/lovely&s.htm

- Fowler C A (1986): 'An event approach to the study of speech perception from a direct-realist perspective', J of Phonetics 14(1):3-28.

- Gibson, J. J. (1966) The Senses Considered as Perceptual Systems. Houghton Mifflin Company,Boston.

- Gibson, J. J. (1979) Ecological Approach to Visual Perception.: Lawrence Erlbaum Associates Publishers, Hillsdate.

- Gregory, R.L. 1988. Consciousness in science and philosophy: conscience and con-science. Chapter 12 in Consciousness in Contemporary Science. (Editors: Marcel, A.J. and Bisiach, E.). Oxford Science Publications.

- Le Morvan, Pierre (2004). Arguments against direct realism and how to counter them. *The American Philosophical Quarterly*, *41*(3), 221-234.] (pdf) http://www.tcnj.edu/~lemorvan/DR_web.pdf

- Oliveira, AndrÃ© L. G. and Oliveira, Luis F. (2002) Toward an ecological conception of timbre. In Proceedings Auditory Perception Cognition and Action Meeting 2002, Kansas City. http://cogprints.org/2713/

- Skinner, B. F. Science and Human Behavior . New York: Macmillan, 1953.

- Skinner, B. F. 1971. Beyond Freedom and Dignity. New York: Knopf.

- Skinner, B. F. 1948. Walden Two. New York: Macmillan.

- Tye, M. (2004). ANOTHER LOOK AT REPRESENTATIONALISM ABOUT PAIN. Consciousness and Emotion 2004, special issue on pain, edited by Murat Aydede (with

replies by M. Aydede, N. Block, B. Maund, and P. Noordhof
http://sun.soci.niu.edu/~phildept/MT/RepresentationalismAndPain.pdf

- Vygotsky, L.S.(1925) Consciousness as a problem in the psychology of behavior. Undiscovered Vygotsky: Etudes on the pre-history of cultural-historical psychology (European Studies in the History of Science and Ideas. Vol. 8), pp. 251-281. Peter Lang Publishing. http://www.marxists.org/archive/vygotsky/works/1925/consciousness.htm

- Ziff, Paul. "About Behaviourism." Analysis 18 (1958): 132-136. Quoted by Larry Hauser in the Internet Encyclopedia of Philosophy. http://www.utm.edu/research/iep/b/behavior.htm#B. F. Skinner: Radical Behaviorism

Indirect Realism

- Moutoussis, K. and Zeki, S. (1997). A direct demonstration of perceptual asynchrony in vision. Proceedings of the Royal Society of London, Series B: Biological Sciences, 264, 393 - 399.

- Rizzo, M., Nawrot, M. & Zihl, J. (1995). Motion and shape perception in cerebral akinetopsia.Brain, 118. 1105-1127. http://nawrot.psych.ndsu.nodak.edu/Publications/Nawrot.pdf/RizzoNawrotZihl.1995.pdf

- Russell, B. (1912). Problems of Philosophy. Home University Library. http://www.ditext.com/russell/rus1.html

- Tye, M. (2003) Qualia. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/qualia/#Repqualia

Further reading:

- Hume, D. (1739). A Treatise of Human Nature. http://www.class.uidaho.edu/mickelsen/ToC/hume%20treatise%20ToC.htm

- Firth, R. (1949) Sense Data and the Percept Theory, Mind, 58 (1949); 59 (1950). (Describes early reasons for rejecting Indirect Realism) http://www.ditext.com/firth/spt.html

- Grush, R. (2000). Self, world and space: The meaning and mechanisms of ego- and allocentric spatial representation. Brain and Mind 1(1):59-92.

http://mind.ucsd.edu/papers/sws/sws.html

The philosophical problem

Chalmers (1996) encapsulated the philosophical problem of consciousness, describing it as the *Hard Problem*. The Hard Problem can be concisely defined as "how to explain a state of consciousness in terms of its neurological basis" Block (2004). A state is an arrangement of things in space over a period of time. It is possible that the Hard Problem has not been solved because the concepts of "space", "time" and "things" are intensely problematic in both science and philosophy.

Some philosophers have argued that *changes* in state are equivalent to "mental states". That consciousness experience always involves acts, such as acts of acquaintance (Russell 1912). But what is a succession of states in the brain or the physical world?

As an extension of the idea of "acts" as mental states many philosophers have argued that the functional description of a system does not need to contain any reference to qualia within that system. Such ideas, based on nineteenth century materialism, have been expressed by Huxley, Ryle, Smart, Goldman and many others. However, although qualia are not required for classical functions, such as most computations or servo-control, it is far from clear whether this is true for all functions. If a function is described as any thing that mediates a change in state it should be realised that "change" itself is not fully understood in philosophy or science and that some systems, such as quantum mechanical systems, contain state changes that are far from understood. It will be seen below that our scientific knowledge is not yet sufficiently complete to allow the claim that all, or even any, changes can occur without qualia.

Whether a philosopher or scientist is dualist, materialist or physicalist they should have some insight into current theories about the physical world. Certainly, if they are considering the problem of "how to explain a state of consciousness in terms of its neurological basis" then some idea of a "neurological basis" is essential.

The objective of this section is to give an account of the problems of space, time and content and to describe how these affect the problem of consciousness.

Epiphenomenalism and the problem of change

Philosophers have noticed since the time of Leibniz that phenomenal consciousness does not seem to be required for the brain to produce action in the Newtonian model of science. Simple explanations of how stimuli at the sense organs might create a signal in the nerves which would be processed by the brain and then create a motion in the muscles do not seem to require phenomenal consciousness. T.H. Huxley is often regarded as the originator of the term **epiphenomenalism** to describe how consciousness seems extraneous to processes in the materialist interpretation of the world although the term may have originated in James' description of Huxley's (1874) ideas.

According to nineteenth century science changes in state cannot explain phenomenal consciousness. It may come as a shock to the reader to discover that nineteenth century science is also unable to account for any change in state. In the materialist paradigm time is construed to be a succession of instants of no duration, each of which is entirely separate from the others. As a result no instant can cause a change in another instant.

On the one hand it seems that conscious experience is not required for a nineteenth century model of behaviour and on the other hand nineteenth century science seems to be impossible without extraneous input from a conscious observer who contains the idea of change.

The problem of change is closely related to the problem of time, which is discussed in depth below. The reader might consider whether phenomenal consciousness is indeed epiphenomenal. Empirical reports describe it as something that is different from the world beyond the body (see direct realism) - but could we generate empirical reports of an epiphenomenon? If we do indeed generate empirical reports of phenomenal consciousness is there some non-materialist, physical** connection between phenomenal consciousness and the functional state?

In the analysis that follows it is essential that the reader does not dismiss the possibility that conscious experience is largely non-functional. The idea that observation is not action should not be dismissed out of hand.

Recommended reading: Mortensen, C. (2002) Change. The Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/change/

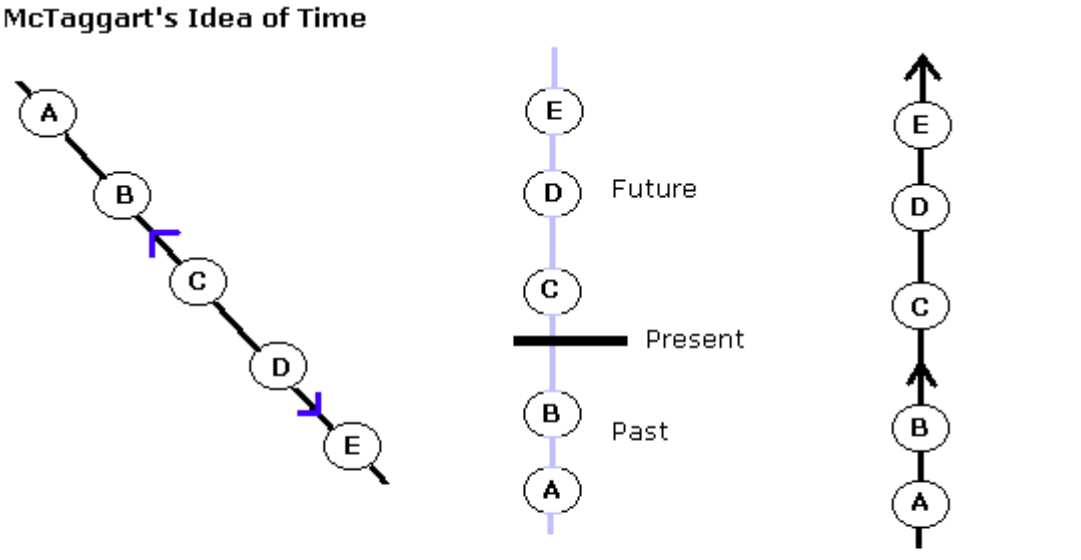
(**) cf: gravity may affect the rate at which clocks tick without the occurrence of any collisions between particles or anything that can be called a "process".

The problem of time

This section should be read after reading a quick introduction to special relativity

The past century of ideas about time

McTaggart in 1908 set out some of the problems with our idea of time in his classic paper *The Unreality of Time*. He drew attention to the way that a sequence of things in a list does not describe time because a sequence of things is constant yet events are always changing. These considerations led him to propose that there are three different sequences of things, or series, that are commonly used to describe events. McTaggart's three different time series are summarised in the illustration below.



He argued that only the 'A Series' is a temporal series because it is only in the A Series that change occurs so that events can be given the labels 'future', 'present' and 'past'. He pointed out that although the A Series is used for determining the direction and sequence of events it is not itself 'in time' because it contains relations that are neither a part of the C Series nor the B Series. This led him to propose that time is unreal because change involves a movement along the time series so cannot be fixed within it.

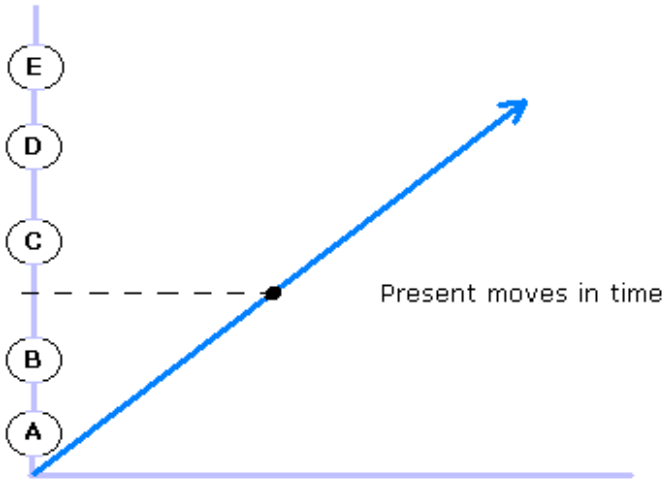
Franck (1994) argued on the basis of Atmanspacher's models of universes with real and imaginary geometries that McTaggart's 'unreality' of time could be avoided by proposing a second, imaginary, time dimension.

"What McTaggart in fact demonstrates is that it is impossible to account for temporality within a strictly one-dimensional concept of time."(Franck 1994).

This idea is illustrated below:

The need for two time axes

McTaggart spotted that the A Series involved relations that were outside of the C or B series and hence not in 'time'. Expressed physically, the A Series is a process in which events move past a cursor (the present):



The C Series, according to McTaggart, is fixed so no motion or process is possible within it. The physical or mathematical interpretation would require another degree of freedom for the A Series to occur - like another time axis.

The idea of time being two dimensional is not new and has also been advanced by such luminaries as Hermann Weyl and CD Broad. Weyl (1920) made the following statement that is extremely apposite to consciousness studies, he wrote that reality is a:

"...four-dimensional continuum which is neither 'time' nor 'space'. Only the consciousness that passes on in one portion of this world experiences the detached piece which comes to meet it and passes behind it, as history, that is, as a process that is going forward in time and takes place in space." (Weyl 1920).

McTaggart's objection to time is felt intuitively by anyone who has contemplated the *Block Universe* of Relativity Theory. If the universe is four dimensional with three space dimensions and one time dimension it would be fixed forever and the observer would be frozen within it. This would occur whether the time dimension was arranged according to Galilean Relativity or Modern Relativity.

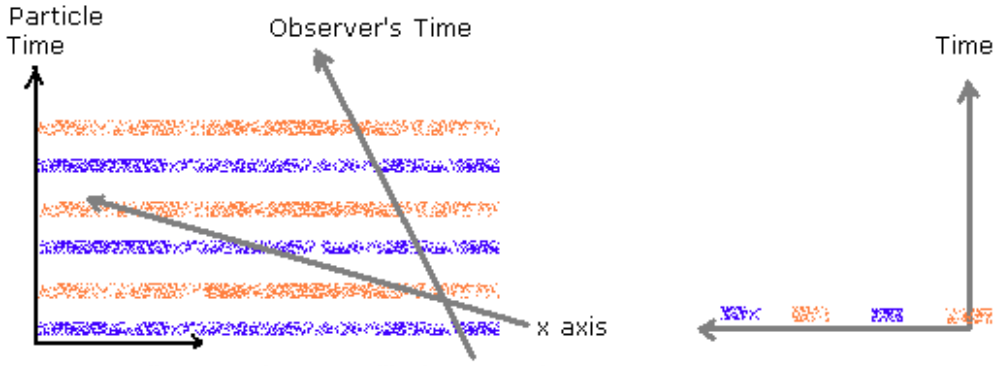
Peter Lynds in 2003 has drawn attention to the 'frozen' nature of the observer in a four dimensional universe. He proposes, like Kevin Brown in his popular mathpages, that time must be approached from the viewpoint of quantum physics because simple four dimensional universes would give rise to 'frozen, static' instants and hence no change could occur. Lynds argues that if quantum physics is introduced then no event can have a definite moment of occurrence and that change occurs because of this quantum indeterminacy:

I would suggest that there is possibly much more to be gleaned from the connection between quantum physics and the inherent need for physical continuity, and even go as far to speculate that the dependent relationship may be the underlying explanation for quantum jumping and with static indivisible mathematical time values directly related to the process of quantum collapse. Time will tell."(Lynds 2003).

Our knowledge of quantum uncertainty can be traced back to De Broglie's highly successful model of individual particle motions. This model was based on Special Relativity theory and it predicted a wave nature for particles. The Heisenberg Uncertainty Principle can be shown to be a consequence of this wave nature. See the illustration below:

De Broglie waves

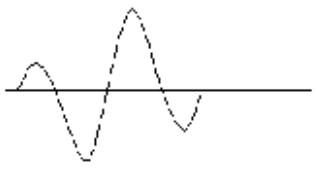
De Broglie waves result from the differing planes of simultaneity for an observer and a moving particle field.



A De Broglie wave is an oscillation occuring simultaneously in the rest frame of the particle but not simultaneously in the frame of the observer.

The observer sees the oscillation as waves distributed in space

Note: the plane of simultaneity of the observer is shown schematically above.



The effect is to produce a spatially distributed pulse

The Heisenberg uncertainty principle can be derived from the fourier transfrom of the pulse ie:

$$\psi(x) = \int_0^{\infty} g(k)\cos(kx)\,dk$$

$$\Delta k \approx 1/(2\Delta x)$$

$$\Delta k\,\Delta x \approx 1/2, \text{ but } k = 2\pi/\lambda = 2\pi p/h = p/\hbar,$$

$$\Delta p\,\Delta x \approx \hbar/2$$

The illustration is based on de Broglie (1925) and Pollock (2004).

So Lynds' argument that change is due to the uncertainty principle is actually an argument that change is due to differing planes of simultaneity between systems that are in relative motion. Kevin Brown is aware of this; he summarises the effect of uncertainty due to special relativity and points out that it provides a resolution of Zeno's arrow paradox:

> "The theory of special relativity answers Zeno's concern over the lack of an instantaneous difference between a moving and a non-moving arrow by positing a fundamental re-structuring the basic way in which space and time fit together, such that there really is an instantaneous difference between a moving and a non-moving object, insofar as it makes sense to speak of "an instant" of a physical system with mutually moving elements. Objects in relative motion have different planes of simultaneity, with all the familiar relativistic consequences, so not only does a moving object look different to the world, but the world looks different to a moving object." (Brown 19??)

Another approach to the way that time has a direction is to suggest that the possible outcomes in quantum mechanics are located in "disjoint space-time regions which exclude one another" (McCall 2000). This does not explain the A Series however because the observer would not have any sense of 'becoming' or temporality as a result of the existence of regions that could not be observed.

Presentism and Four-Dimensionalism

In the past century the philosophical battle lines have been drawn between the Presentists, who believe that only the durationless instant of the present exists and the Four Dimensionalists who consider that things are extended in both space and time (see Rea (2004)). There are two types of Presentism, in its extreme form it is the belief that the past and future are truly non-existent, that what we call time is not an axis for arranging things but a series of changes and records in an *enduring* present. In its less extreme form, which might be called *functional presentism*, the present is a durationless instant that can never be connected to the future or past except through predictions and records.

In consciousness studies it is the conventional theory that brain activity occurs in the present instant and that the past can only occur as memories retrieved into this durationless present. So, in consciousness studies functional Presentism seems to be the accepted paradigm.

Presentism cannot explain change. Each instant is durationless and frozen. That said, as seen above, four dimensionalism cannot explain the observation of change although it can explain the difference between moving and stationary objects. Fortunately the debate has been largely resolved by recent scientific experiments which show that time exists and hence Presentism is unlikely.

The existence of time

The issue of whether or not time exists is critical to consciousness studies. If we exist at an instant without duration then how can we know we exist? Clay (1882) coined the term 'specious present' to describe how we seem to exist for a short period containing the immediate past:

"All the notes of a bar of a song seem to the listener to be contained in the present. All the

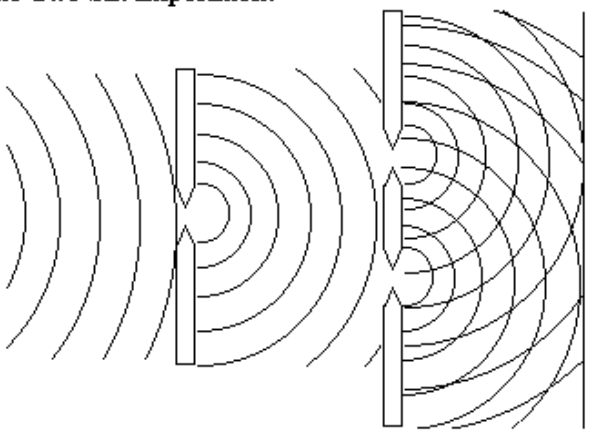
changes of place of a meteor seem to the beholder to be contained in the present. At the instant of the termination of such series, no part of the time measured by them seems to be a past. Time, then, considered relatively to human apprehension, consists of four parts, viz., the obvious past, the specious present, the real present, and the future."

So conscious, phenomenal experience has things that are apparently extended in time. But does time exist?

Recent experiments in quantum physics should change our view of time forever. Lindner et al (2005) have explored the problem of time by investigating quantum interference between interferometer slits that are separated by time rather than space.

In the famous, spatial 'double slit experiment' in quantum physics single electrons are directed at an apparatus that has the equivalent of two tiny slits separated by a small gap. The electrons pass through the apparatus one at a time and produce flashes of light on a screen or changes in a photographic plate. The electrons produce series of bands on the screen that are typical of interference effects. So each electron is deflected as if it has passed through both slits and interfered with itself.

The Two Slit Experiment



Single slit Double slit Screen

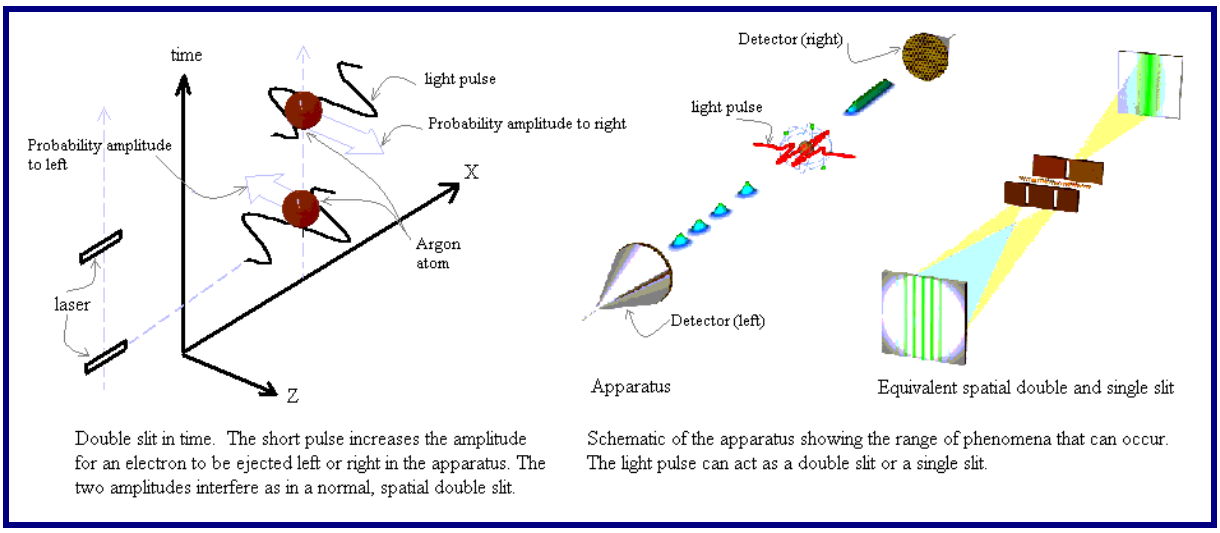
It is found that electrons make a full interference pattern on the screen

Water waves, light waves etc all produce interference patterns if they are passed through two slits. Each slit acts as a new source of waves. If electrons are passed through a two slit apparatus ONE AT A TIME an interference pattern appears on the screen after a few thousand electrons have passed through.

This experiment provided some of the earliest evidence for the wave-packet nature of the electron.

In an amazing technical tour de force Lindner et al (2005) have extended the idea of the spatial double slit experiment to an investigation of time. In the double slit experiment in time electrons are produced in an inert gas by extremely short laser pulses. The pulses stimulate a single atom and there is a probability of this atom releasing an electron at each oscillation of the pulse. The apparatus is described by Paulus et al (2003). The probability (see note 1) of an electron being ejected to the left or right of the apparatus can be adjusted by adjusting the optical pulse. Pulses can be applied with a duration of a few femtoseconds and these create 'slits' extending over an interval of about 500 attoseconds (500 x 10-18

seconds). A single electron has a probability of being emitted at each of the slits. The probability of the single electron going in a particular direction after both slits have been created depends upon the interaction of the probabilities of being emitted in a particular direction at each single slit. As expected, an interference pattern was generated as a result of single electrons interfering with themselves across different times.



Double slit in time. The short pulse increases the amplitude for an electron to be ejected left or right in the apparatus. The two amplitudes interfere as in a normal, spatial double slit.

Schematic of the apparatus showing the range of phenomena that can occur. The light pulse can act as a double slit or a single slit.

This experiment is remarkable because it provides direct evidence that time exists in a similar fashion to the way that space exists. It is consistent with Feynman's theory of Quantum Electrodynamics where all possible paths, both in time and space, interact to produce the final trajectory of a particle and consistent with modern Special Relativity, on which QED is based, where the trajectories of particles occur in an extended four dimensional space-time.

The experiment has not attracted as much attention as it might have done because most physicists are not Presentists. To physicists the experiment is yet another confirmation of modern physics. However it has impressed many:

"This experiment should be included in every textbook on quantum mechanics," says Wolfgang Schleich, a quantum physicist at the University of Ulm in Germany. "It certainly will be in mine." (PhysicsWeb)

Why should a concrete demonstration that time exists affect consciousness studies? The simple answer is that, as Kant, Gombrich, Clay, James and many others have spotted, there can be no conscious, phenomenal experience without time. The fact that time exists should provide new insights and liberate theorists in the field of consciousness studies from the problems of recursion and regression that are inherent in Presentism.

Meanwhile Quantum Theorists are pressing on with the problem of how an organised spacetime could emerge from quantum chaos (cf: Ambjorn et al (2004)) and even how mind might be involved in the emergence of time itself (cf: Romer (2004)).

The nature of time

The nature of classical time

In the eighteenth century it became apparent that Euclid's parallel postulate could not be explained in terms of the other postulates. The parallel postulate is equivalent to the statement that exactly one line can be drawn through any point not on a given line in such a way that it is parallel to the given line (this is Playfair's simple version). It is also known as the fifth postulate.

The attempts to prove the parallel postulate led to the development of non-Euclidean geometry. It was then possible to show that the parallel postulate is a special case within a range of geometrical forms from spherical geometry, through Euclidean geometry to the hyperbolic geometry of Bolyai and Lobatschefsky. Furthermore it was shown by Taurinus that the axioms of Euclidean geometry, with the exception of the fifth postulate, applied on the surface of a shere with an imaginary radius. This motivated Hermann Minkowski to propose that Einstein's new theory of relativity was in fact due to the universe being a 'space-time' with four dimensions rather than just a space in which things change (see Walter 1999). In 1909 Minkowski said that:

> "Henceforth space by itself and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality". (Minkowski 1909).

The earliest idea of the four dimensional universe involved time as an axis with displacements measured in units of the square root of minus one (cf: Einstein (1920)): time was considered to be displacements along the imaginary plane. However, from the moment of Minkowski's proposal mathematicians were aware that other interpretations of time could give almost identical physical results.

According to the differential geometry developed during the nineteenth century a space is defined in terms of a *metric tensor* which is a matrix of factors that determine how displacements in each independent direction vary with displacements in the other directions. The metric tensor then specifies a *metric* which is an equation that describes the length of a displacement in any direction in terms of the independent directions, or *dimensions*.

A derivation of the metric tensor and how it can be used to calculate the metric is given in Consciousness studies:The philosophical problem - Appendix.

The metric of the space considered by Euclid is Pythagoras' theorem where the length of any displacement is given in terms of the displacements along the three independent axes, or dimensions:

$$s^2 = x^2 + y^2 + z^2$$

It is interesting to explore *imaginary time* from the point of view of consciousness studies. Minkowski's original idea for the geometry of the world proposed that any displacement was a displacement in both time and space given by a four dimensional version of Pythagoras' theorem:

$$s^2 = x^2 + y^2 + z^2 + (ict)^2$$

which, given that $i^2 = -1$ equals:

$s^2 = x^2 + y^2 + z^2 - (ct)^2$

Where *i* is the square root of minus one, *c* is a constant for converting metres to seconds and t is the displacement in time. The space-time is considered to be flat and all displacements are measured from the origin.

The interesting feature of Minkowski space-time with imaginary time is that displacements in time can *subtract* from displacements in space.

If we set $r^2 = x^2 + y^2 + z^2$ (where *r* is the radius of a sphere around the origin then:

$s^2 = r^2 - (ct)^2$

Notice that $s^2 = 0$ when $r^2 = (ct)^2$ so if imaginary time existed there would be times and separations within a spherical volume of things where **everything is at a point as well as distributed in space**. This idea has distinct similarites with the *res cogitans* mentioned by Descartes, and the *point soul* of Reid and Malebranche etc., however, this feature of Minkowski's space-time has not been popular with physicists for some good reasons. Blandford and Thorne point out some of the problems:

One approach, often used in elementary textbooks [and also used in Goldstein's (1980) Classical Mechanics and in the frst edition of Jackson's Classical Electrodynamics], is to set $x^0 = it$, where and correspondingly make the time basis vector be imaginary,... When this approach is adopted, the resulting formalism does not care whether indices are placed up or down; one can place them wherever one's stomach or liver dictate without asking one's brain. However, this $x^0 = it$ approach has severe disadvantages: (i) it hides the true physical geometry of Minkowski spacetime, (ii) it cannot be extended in any reasonable manner to non-orthonormal bases in flat spacetime, and (iii) it cannot be extended in any reasonable manner to the curvilinear coordinates that one must use in general relativity. For this reason, most advanced texts [including the second and third editions of Jackson (1999)] and all general relativity texts take an alternative approach, which we also adopt in this book. This alternative approach requires introducing two different types of components for vectors, and analogously for tensors: contravariant components denoted by superscripts, and covariant components denoted by subscripts." Blandford & Thorne (2004).

What Blandford and Thorne are saying is that the metric of space-time appears to be the result of the interaction of two coordinate systems and cannot be explained by a single coordinate system with imaginary time. When a more complicated geometrical analysis is applied it is evident that there are two possibilities for the time coordinate. In the first the metric can be **assumed** from the outset to be

$s^2 = x^2 + y^2 + z^2 - (ct)^2$

and the metric tensor simply adjusted by inserting -1 in the principle diagonal so that the negative sign in front of the time coordinate occurs. With this assumption and adjustment the time coordinate can be assumed to be *real*. In the second possibility the time coordinate in the world can be assumed to be imaginary and the time coordinate of the observer can be assumed to be real. This gives rise to the same metric tensor and metric as the first possibility but does not assume the resulting metric from the outset.

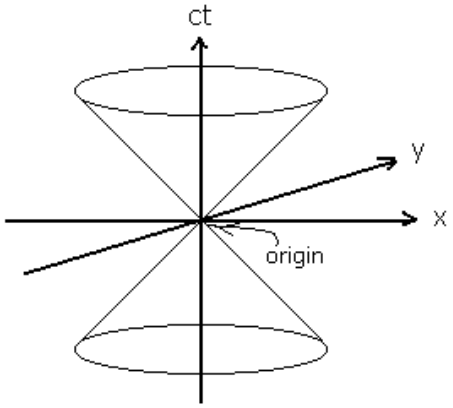
The three ideas of classical time (imaginary, real and mixed) are shown in the illustration below:

The light cone.

The light cone is given by the equation:

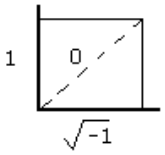
$0 = x^2 + y^2 + z^2 - (ct)^2$

(the z dimension is suppressed in the diagram to allow the equation to be shown on paper).



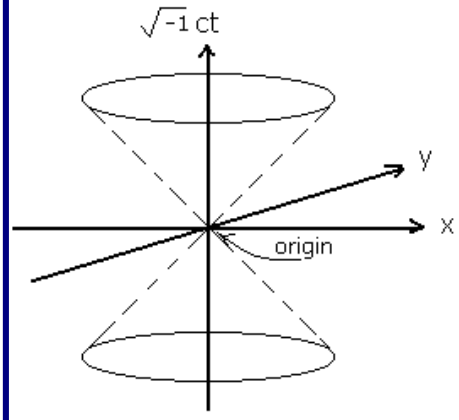
An interesting feature of imaginary units is the way they can subtract from real displacements.

$0 = (1)^2 + (\sqrt{-1})^2$



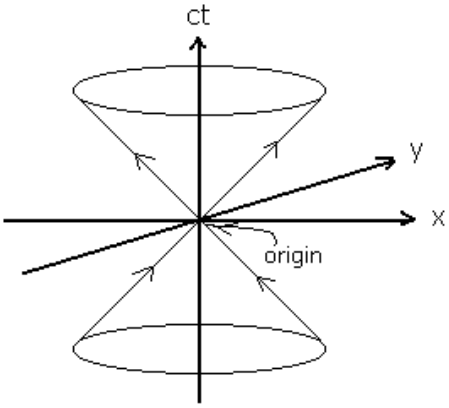
Imaginary Time

If time is imaginary then the surface of the light cone is literally in contact with the origin.



Real Time

If time is real then the surface of the backward light cone contains the path of light rays into the origin and the surface of the forward light cone contains the path of light rays away from the origin



Mixed real and imaginary time

Opens up possibilities such as the observed world behaving as if there were only real time but the universe as a whole operating with both types of time.

The light cone is divided into three regions: events on the surface of the light cone, such as photons converging on the observer, are said to be *lightlike* separated from the observer, events inside the future or past light cones are said to be *timelike separated* and events outside the lightcone are said to be *spacelike* separated from the observer.

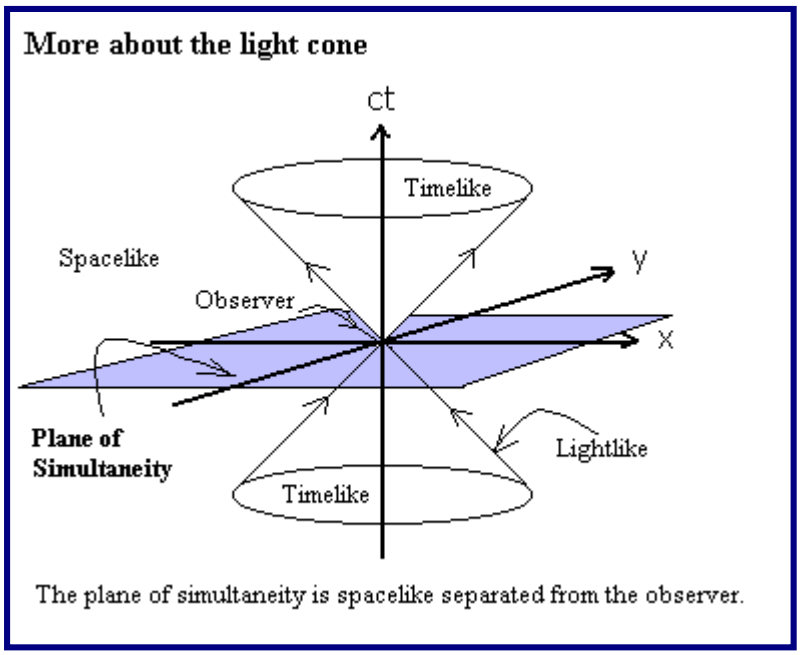
The physical **theory of relativity** consists of four dimensional geometry plus the assumption of causality and the assumption that physical laws are invariant between observers. It should be noted that space-time could contain preferred frames of reference and is not, by itself, a theory of relativity. The assumption that physical laws are invariant between observers leads to the postulate that nothing can travel faster than *c* metres per

second. This means that the constant *c*, which in Minkowski space-time is the conversion factor from seconds to metres then has a new significance as the maximum velocity.

A result of *c* being a maximum velocity is that nothing can travel from regions of the light cone that are spacelike separated to the observer at coordinates (0,0,0,0). This is problematic for observers if time is real because, as Stein (1968) wrote:

"in Einstein-Minkowski space-time an event's present is constituted by itself alone." (Stein 1968).

However, to each of us it seems that the present is characterised by *many* things simultaneously. As will be discussed below, this simultaneity of present things also results in the appearance of phenomenal space. But in Minkowski space-time with real time the plane of simultaneity is entirely space-like separated from the observation point. If real time is accepted it would appear that we cannot have the space of phenomenal experience. The regions of the light-cone and the spacelike separation of present events are shown in the illustration below:



So can the time in Minkowski space-time be real? If time were in some way related to the imaginary plane then all the content of the surface of the light cone could be simultaneously at the position of the observer and phenomenal experience containing space is possible, but then general relativity may be problematic. So can the time in Minkowski space-time be imaginary?

There is another problem with Minkowski space-time known as the "Rietdijk-Putnam-Penrose" argument or the Andromeda paradox (Penrose 1989). Moving observers have different planes of simultaneity. The plane of simultaneity of an observer moving towards you slopes upward relative to your plane of simultaneity (see the illustration on "De Broglie waves" above). Suppose an alien civilisation in the Andromeda galaxy decided to launch a fleet of spacecraft intent on the invasion of earth just as you passed Jim in your car. Your plane of simultaneity would slope upwards ever so slightly compared with Jim's, Jim's plane of simultaneity could contain earlier events on Andromeda than yours. At the distance of the

Andromeda galaxy it could be another week or two for the Andromedean's to launch their invasion fleet in Jim's slice of the universe. Penrose considers that this example shows that the events in the universe must be fixed:

"Two people pass each other on the street; and according to one of the two people, an Andromedean space fleet has already set off on its journey, while to the other, the decision as to whether or not the journey will actually take place has not yet been made. How can there still be some uncertainty as to the outcome of that decision? If to either person the decision has already been made, then surely there cannot be any uncertainty. The launching of the space fleet is an inevitability." (Penrose 1989).

If the decision to invade and a time previous to this decision are both part of the present instant on earth then, in a 4D classical universe, the decision to invade must be inevitable. This lack of free will in a 4D universe is known as chronogeometrical determinism (Toretti 1983). However, as de Broglie demonstrated, it is sloping planes of simultaneity that do indeed introduce uncertainty into our universe. It should also be noted that nothing on the plane of simultaneity is observable to the owner of that plane because, to observe it would involve the transmission of data at velocities greater than the speed of light.

Petkov (2002)considers a version of the Andromeda paradox in depth. He concludes that:

"If the relativity of simultaneity is explicitly discussed in terms of the dimensionality of reality, the fact that observers in relative motion have different sets of simultaneous events can be explained either by assuming that existence is also relativized (preserving the views of the present and objective becoming) or by considering existence absolute which means that reality is a 4D world. Although the option of relativizing existence appears completely unacceptable from a philosophical point of view, that option is eliminated within the framework of SR by demonstrating that the twin paradox would not be possible if existence were not absolute."

According to Petkov Special Relativity describes the universe as a frozen space-time where things are eternally arranged in four dimensions. Petkov introduces the possibility of change as a feature of consciousness and in support of this quotes Weyl's intuition that only the conscious observer moves in time.

Relationalism, Substantivalism, the Hole Argument and General Covariance

Relationalism and Substantivalism

The view that the universe could be an extended space and time with things in it, a sort of unbounded container, is known as **substantivalism**. It was championed by Newton and Clarke in the seventeenth century. The view that the space and time in the universe depends upon the relations between the objects in the universe is known as **relationalism** and was championed by Leibniz.

Leibnitz attacked substantivalism by arguing that if there were two universes which only differed by things in one universe being displaced by five feet compared with things in the other universe then there is no reason why the two universes should be discernably different.

Newton supported substantivalism by arguing that when the water in a bucket rotates it adopts a concave surface that is independent of other motions and provides evidence of the possibility of absolute motion. This argument is called the *bucket argument*. Newton also introduces the *globe argument* in which he proposes that the state of motion of two globes connected by a taut thread can be gauged from the tension in the thread alone. When the globes are stationary with respect to each other there is no tension in the thread.

Ernst Mach in 1893 introduced a relationalist account of the bucket argument by claiming that the water rotates in relation to the fixed stars. He stated this in what has become known as Mach's principle:

"The inertia of any system is the result of the interaction of that system and the rest of the universe. In other words, every particle in the universe ultimately has an effect on every other particle."

The relationalist position is interesting from the viewpoint of consciousness studies because phenomenal consciousness appears as a projection that overlies physical space. As an example, the stars on the ceiling of a planetarium appear to be at huge distances from the observer even though they are reflected lights that are only a few metres away. In general a projection where positions depend upon angular separations will be subject to relationalism. It is also probable that the space of phenomenal consciousness is a continuum of some field in the brain, if this is the case then the way we conceive of space as an existent entity is actually a conception involving the angular relations between the perturbations of the substance that is the field. Substantivalism would then literally be space as a substance. It is intriguing in this respect that Kant believed that space was a form of intuition and hence a property of mind.

Kant raised another type of argument for the justification of absolute space. He asked whether *handedness* was due to relations or a property of space. The right and left hands are enantiomorphs (mirror images). The relations within the right and left hands are identical but they still differ, for instance a right hand cannot be moved on to a left hand so that it exactly overlies it. Kant proposed that handedness was property inherent in space itself rather than a set of relations.

Gardner introduced a version of Kant's problem with the "Ozma" argument:

"Is there any way to communicate the meaning of the word "left" by a language transmitted in the form of pulsating signals? By the terms of the problem we may say anything we please to our listeners, ask them to perform any experiment whatever, with one proviso: there is to be no asymmetric object or structure that we and they can observe in common." (Gardner 1990).

Although it is probably impossible to provide an answer to the Ozma argument it is possible to relate handedness to a conceptual point observer who spans more than an instant of time. If a point observer is at the centre of a field of inward pointing space-time vectors then relative to any given vector there are positive and negative angular separations. The body is asymmetric and the point observer would lie within this so always have available a 'head' direction or a 'foot direction' and hence a left and right. Unlike the time extended observer an instantaneous observer would not contain vectors that contained directional information and would be no more than a collection of points in space.

Pooley (2002) discusses handedness in depth and introduces the problem of parity violation

in the Weak Interaction.

General Covariance and the Hole Argument

The proposal that the universe is four dimensional does not in itself produce a full physical theory. The assumptions of causality and the invariance of physical laws between observers are also required to create modern Relativity Theory. The second assumption, that the laws of physics are the same for all observers is closely related to the requirement of **general covariance**.

The principle of general covariance requires that a manifold of events can be smoothly mapped to another manifold of the same dimension and back again. This mapping should always give the same result. General covariance is assumed in General Relativity.

Einstein realised that there was an apparent problem with this assumption in certain circumstances. In his **hole argument** he considers a special region of space-time that is devoid of matter and where the stress-energy tensor vanishes. He then labels the same events outside the hole with two different coordinate systems. These coordinate systems could differ by something as simple as having origins that are separate so the difference is entirely passive. Both systems will give the same values for the gravitational field outside the hole. It turns out however that that the systems predict different fields within the hole (see MacDonald (2001) for the calculation and Norton (1993), (1999) for a discussion). Einstein overcame this problem by considering active mappings where particles are actually transferred through the hole. He concluded that the points where particles meet can be transformed according to general covariance and hence a relativistic theory could indeed be constructed. Solutions to the field equations that were inconsistent with the points defined by interacting particles were discarded as non-physical.

The hole argument led Einstein to abandon the idea of space and time as something separate from the material content of the universe. The General Theory of Relativity becomes a theory of **observables**. He wrote that:

"That the requirement of general covariance, which takes away from space and time the last remnant of physical objectivity, is a natural one, will be seen from the following reflection. All our space-time verifications invariably amount to a determination of space-time coincidences. If, for example, events consisted merely in the motion of material points, then ultimately nothing would be observable but the meetings of two or more of these points. Moreover, the results of our measurings are nothing but verifications of such meetings of the material points of our measuring instruments with other material points, coincidences between the hands of the a clock and points on the clock dial, and observed point-events happening at the same place at the same time. The introduction of a system of reference serves no other purpose than to facilitate the description of the totality of such coincidences". (Einstein 1916).

This is what would be expected from a four dimensional block universe with real time. It is a frozen universe of the type discussed earlier. As Earman (2002) puts it when discussing change:

"First, the roots of the problem lie in classical GTR, and even if it was decided that it is a mistake to quantize GTR, there would remain the problem of reconciling the frozen dynamics of GTR with the B-series notion of change that is supported not only by common sense but by every physical theory prior to GTR. Second, although the aspect of the

problem that grabs attention is that of time and change, no solution will be forthcoming without tackling the more general issue of what an 'observable' of classical GTR is."

In such a universe action at a distance is not possible. From the viewpoint of consciousness studies the limitation of physical concepts to interactions between particles is a restatement of Ryle's regress and the recursion version of the homunculus problem. If events are no more than space-time coincidences then we are doomed to the endless transfer of data from point to point without any conscious observation. This seems to forbid any true simultaneity in experience and means that only measurements are possible.

The reduction of physics to the study of particle interactions is fully relationalist and allows space-time to become a property of these interactions rather than vice-versa. Once it becomes possible to consider space-time as a dependent property it is then feasible to equate *observation* with *measurement*. Observation is normally the representation of an event in an observer's space-time coordinate system. Measurement is the change in state of a system in response to an encounter with an event. If we maintain that space-time does not exist and can be replaced by encounters between particles then observation can be replaced by measurement. This may well be a way forward for some approximations to physical reality and may allow us to understand how a space-time is selected within an observer. As part of this approach the word "observable" is often used interchangeably with "measurable".

Quantum theory and time

The general problem of QM and time

Quantum physics provides many fundamental insights into the nature of time. At the simplest level the energy-time version of the Heisenberg Uncertainty Principle predicts that Quantum Mechanical (QM) interference should occur between a particle and earlier versions of itself. Such interference has been observed (see "The existence of time" above).

Two of the most complete reviews of the problem of time in quantum theory available at present are Zeh (2001) and Isham (1993).

Perhaps the most interesting aspect of QM and time is that it can provide an argument that time does not exist in the universe as a whole. The argument can be approached from many directions (See Rovelli 2003) but is clear in the Wheeler-de Witt equation which describes the wavefunction of the entire universe. This wavefunction has no reference to time. De Witt explained the emergence of time by proposing that the universe can be divided into an observer with measuring instruments and the rest of the universe so that the rest of the universe changes with respect to the observer.

Rovelli (2003) supports this idea of partition, he considers in depth the problems of the "hole argument" and quantum physics and notes that, given the assumption that events are just successions of relations:

"The unique account of the state of the world of the classical theory is thus shattered into a multiplicity of accounts, one for each possible "observing" physical system. Quantum mechanics is a theory about the physical description of physical systems relative to other systems, and this is a complete description of the world. (Rovelli 2003).

Barbour (1997) and Hartle and Gell-Mann have both proposed that an observer is a partition

or region with memories that contain the trace of histories. The histories would represent a B Series. Unfortunately this leaves the A Series unexplained so time would have a direction but there would be no 'becoming'.

Hawking introduces the observer into the problem of time by asking what sort of universe is compatible with human life. This application of the **Anthropic Principle** leads to constraints on the form of the universe, for instance the universe should have galaxies and last for more than a few million years. The Anthropic Principle is actually a restatement of the observer problem - if being an observer leads to a certain division of the universe into observer and observed then the observed part will have the form given by the Anthropic Principle. Hartle and Hawking () also tackled the "boundary problem" of cosmology by proposing that there is no boundary. This proposal involves adding a fifth, time-like, dimension on the imaginary plane so that the universe at it's beginning is a **de Sitter** or **anti de Sitter** space-time.

A de Sitter space-time is characterised by the metric:

$ds^2 = dx^2 + dy^2 + dz^2 + (idt)^2 + du^2$

An anti de Sitter space time has the metric:

$ds^2 = dx^2 + dy^2 + dz^2 + (idt)^2 - du^2$

A de Sitter space time is fascinating from the view point of consciousness studies because it contains three space-like dimensions, one real, time-like dimension (u) and one imaginary time-like dimension. This might give the real and imaginary time-like axes that Franck proposed were needed to produce the McTaggart A Series. However, the extra dimension could only be related to the observer in the universe as it is at present because the extra dimension does not appear to be required to explain measurables.

The interpretation of QM

Time is also of interest in the interpretation of quantum mechanics and entanglement. There are many interpretations of QM such as the **Operational Interpretation** (Decoherence Theory), the **Transactional Interpretation**, the **Relational Interpretation**, the **Many Worlds Interpretation**, the **Copenhagen Interpretation**, the **Bohm Interpretation**, the **Many Minds Interpretation etc.**

Some of these interpretations, such as the Transactional Interpretation, allow the connection of entangled quantum states backwards in time along the path of particles.

Decoherence theory is of particular interest because it allows the calculation of how long an entangled state can persist. Tegmark (2000) and Hagan et al (2002) have used this technique to calculate the decoherence time of entanglement in microtubules and have differed by a factor of 10^{10} because of differing assumptions about the biophysics of microtubules in the brain.

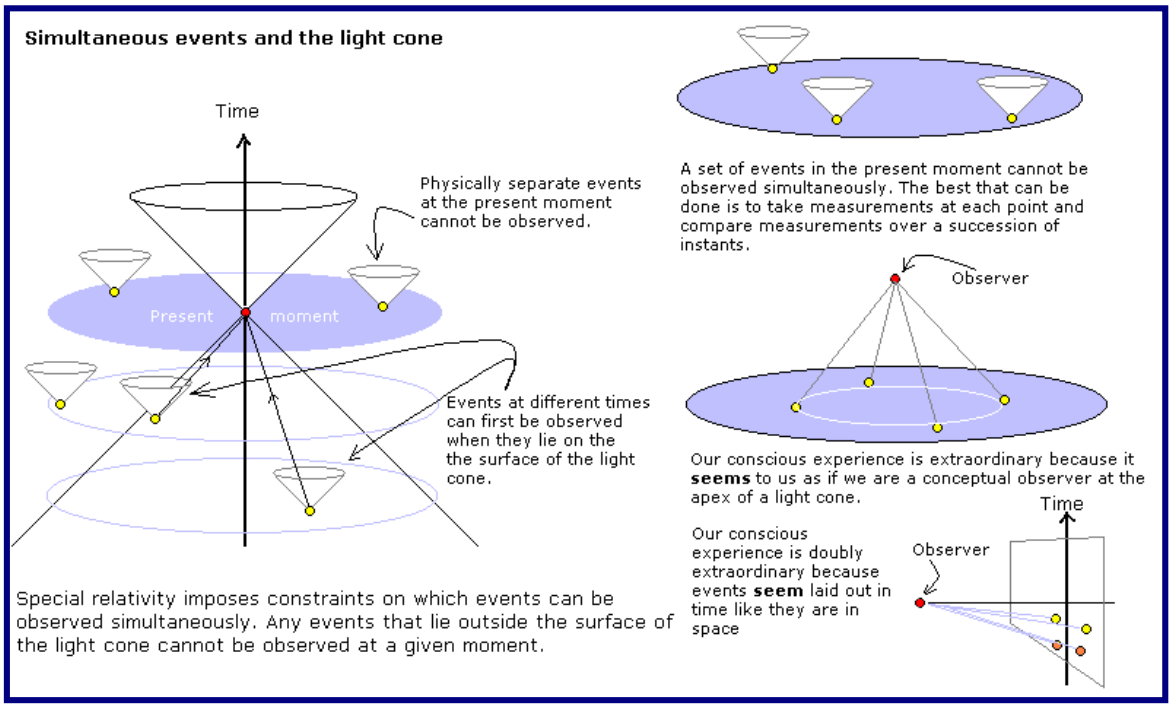
Time and conscious experience

In a four dimensional universe time is an independent direction for arranging things. As an independent direction things arranged in time do not overlie things arranged in space. This

also appears to be the case in conscious experience where whole words or "bars of a tune" can be experienced arranged in time. This extension in time is easy to experience but the independence of the time dimension is difficult to conceive, for instance Le Poidevin (2000) reflects that:

> "If events e1 and e2 are registered in a single specious present, then we perceive them both as present, and so as simultaneous. But we do not see, e.g., the successive positions of a moving object as simultaneous, for if we did we would see a blurred object and not a moving one."

This assumes that arrangements in time do not occur in an independent direction for arranging things and hence would overlay space. In fact the mystery of conscious experience is deeply related to how we can experience many things as events that are separate from each other. Our experience of two dimensional patterns containing many things is as much a mystery as how we experience temporal patterns extended in time. The problem is illustrated below:



It is as if patterns in conscious experience are being viewed from a point in at least four dimensions. How our experience can be like the 'view' of a conceptual point observer who is at the apex of a light cone without the data being overlaid and obscured is a profound mystery, obviously the data cannot be transferred into the apparent observation point and appears as nebulous vectors directed at the point. Some philosophers have noticed this problem.

(*This is a stub, requires an elaboration of Specious Present Theory and Husserl's ideas*)

Le Poidevin (2000). The experience and perception of time. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu//archives/spr2001/entries/time-experience/#4

The problem of space

The problem of Relationalism and Substantivalism has been discussed earlier. In this section the concept of space will be explored in more depth.

Space is apparent to us all. It is the existence of many simultaneous things at an instant. If we see a ship and hear a dog barking on our left there is space. If we look at a checkerboard there is space. This occurrence of space in phenomenal experience is similar to the measurement of space in the world: things that are simultaneously at the ends of a metre rule are a metre apart; if there is more than one object at a given instant the objects are separated by space.

Physicists have found that the mathematics of **vector spaces** describes much of the arrangement of things in the world. In a vector space the independent directions for arranging things are called **dimensions**. At any instant physical space has three clearly observable dimensions.

It has been known for millenia that the three dimensions observable at an instant are interrelated by **Pythagoras' Theorem**:

Pythagoras' theorem on a plane shows that the length of any displacement is related to the sum of the squares of the displacements in the independent directions (x and y):

$h^2 = x^2 + y^2$

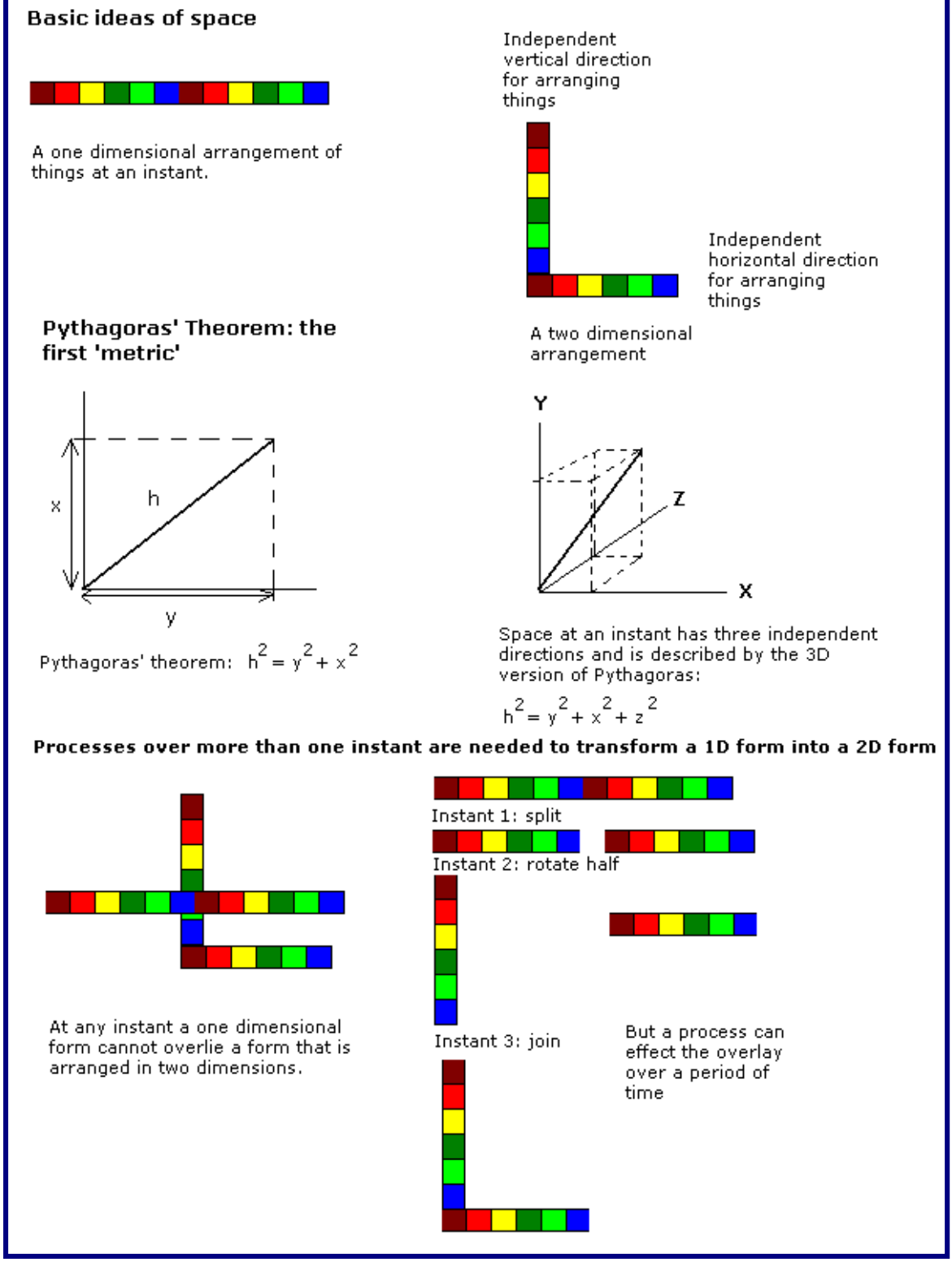
Pythagoras' theorem in three dimensions is:

$h^2 = x^2 + y^2 + z^2$

The advances in geometry in the nineteenth century showed that Pythagoras' theorem was a special case of a **metric**, an equation that describes displacements in terms of the dimensions available. In the twentieth century it was realised that time was another independent direction for arranging things that was interrelated to the other three dimensions. The world is now described as a **four dimensional manifold**.

The illustration overleaf shows how different numbers of dimensions affect the arrangement of things.

It is sometimes suggested that our idea of space is due to some sort of memory that is read out sequentially. This is unlikely because, at any instant a one dimensional form cannot be made to overlie a two dimensional form and a two dimensional form cannot overlie a three dimensional form etc. One dimensional forms are not *congruent* with two dimensional forms. This means that a one dimensional form such as *virtual memory* cannot, at any instant, overlie two dimensional forms such as occur in phenomenal experience and hence experience does not supervene on the idea of virtual memory (See section on functionalism as a one dimensional Turing Machine).

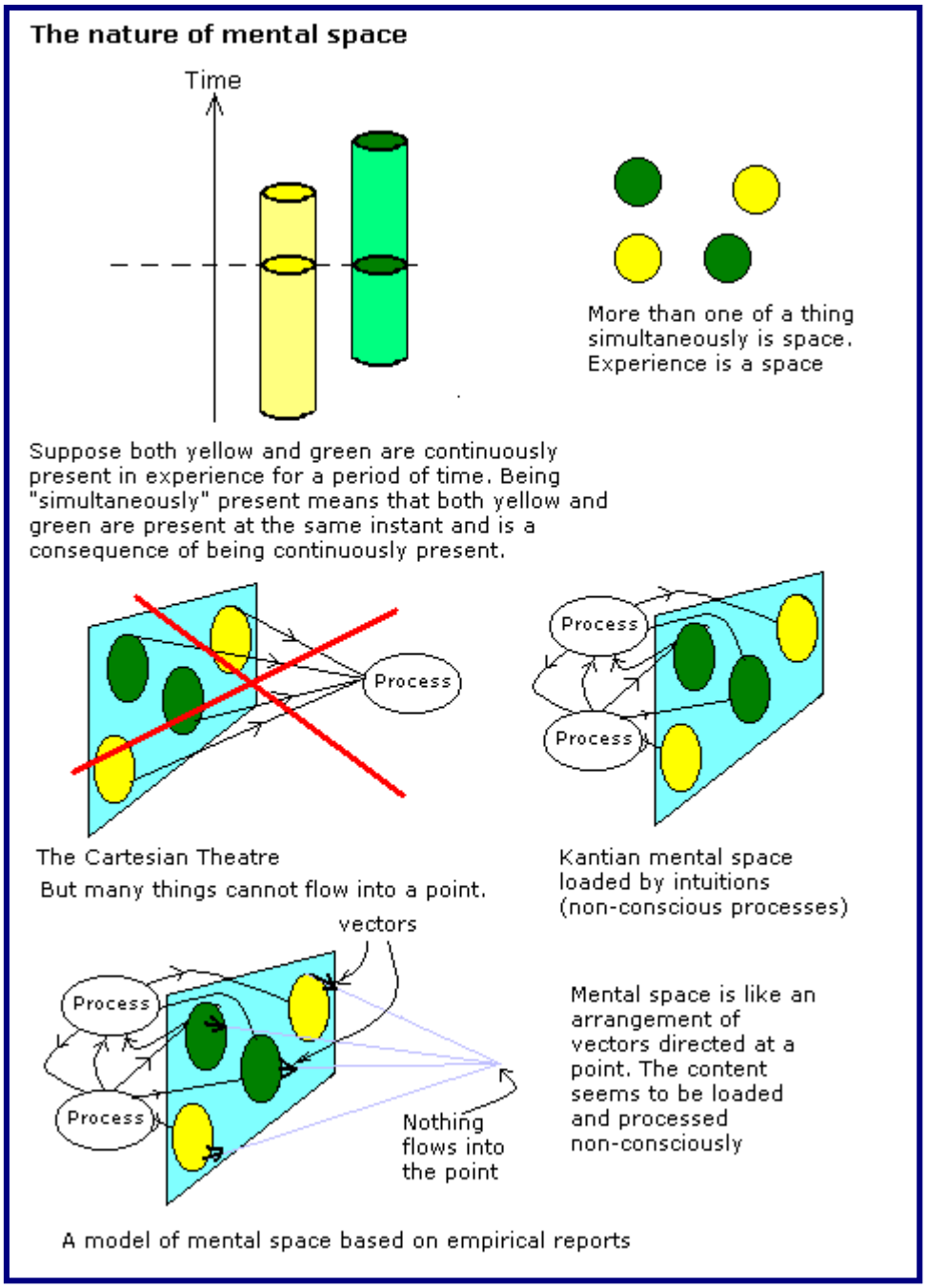


Curiously the idea of *mental space* is often denied. McGinn(1995) gives such a denial:

"We perceive, by our various sense organs, a variety of material objects laid out in space, taking up certain volumes and separated by certain distances. We thus conceive of these perceptual objects as spatial entities; perception informs us directly of their spatiality. But conscious subjects and their mental states are not in this way perceptual objects. We do not see or hear or smell or touch them, and a fortiori do not perceive them as spatially individuated.(2) This holds both for the first- and third-person perspectives. Since we do not observe our own states of consciousness, nor those of others, we do not apprehend these

states as spatial." McGinn(1995).

This denial is strange because it begins by describing phenomenal experience as clearly spatial and then proceeds to argue that there is some other thing, the "mental state", which is non-spatial. This seems to contradict our everday life where our experience is our experience, there is no other experience.



The issue is whether this experience is things in themselves (Direct Realism) or some other form in the brain (Indirect Realism). The illustration above shows how space occurs in phenomenal experience; it sidesteps the issue of the location of the contents of phenomenal

consciousness.

McGinn (1995) gives a description of how phenomenal experience cannot be overlaid by a 3D model of events in the brain:

"Consider a visual experience, E, as of a yellow flash. Associated with E in the cortex is a complex of neural structures and events, N, which does admit of spatial description. N occurs, say, an inch from the back of the head; it extends over some specific area of the cortex; it has some kind of configuration or contour; it is composed of spatial parts that aggregate into a structured whole; it exists in three spatial dimensions; it excludes other neural complexes from its spatial location. N is a regular denizen of space, as much as any other physical entity. But E seems not to have any of these spatial characteristics: it is not located at any specific place; it takes up no particular volume of space; it has no shape; it is not made up of spatially distributed parts; it has no spatial dimensionality; it is not solid. Even to ask for its spatial properties is to commit some sort of category mistake, analogous to asking for the spatial properties of numbers. E seems not to be the kind of thing that falls under spatial predicates. It falls under temporal predicates... McGinn(1995)

He concludes that a 3D form can only be rearranged into the form of the things in experience over a succession of instants ("It falls under temporal predicates"). This is highly suggestive of phenomenal experience having more than three dimensions in the same way as an ordinary physical thing or field has more than three dimensions.

The problem of qualia

A quality of an object such as its colour, roughness, temperature etc. is known as a **quale**, the plural of quale is **qualia**. Qualia are the contents of phenomenal consciousness.

The physics of qualia

According to physicalism qualia must be things in the universe. But what are "things in the universe" and which of these are qualia?

If we wish to explain phenomenal experience we must first decide whether experience is a measurement or things themselves. Measurement begins with a quantum mechanical interaction between an instrument and a set of particles, this then creates a *signal* which is a change in the state of the instrument. The signal can be a flow of charge or a chemical change etc. In the Direct Realist case the signal would be the change at the interface between the bulk of a material (a crude measuring device) and a set of QM particles, in the Indirect Realist case it would be some signal in the brain derived from the initial signal. In either case phenomenal consciousness would be some form of a set of signals themselves. Intriguingly, these final signals that are the content of consciousness would be subject to quantum uncertainty.

The signals that form phenomenal consciousness would differ from those that mediate the transfer of information from QM phenomena to measuring instruments.

The signals in measuring events arise as a result of interactions between QM phenomena and a measuring apparatus composed of relatively large structures in the environment. These structures (called *the environment*) produce signals at definite locations. This chain of fixing positions is known as **decoherence** (see Zurek (2003) or Bacciagaluppi (2004) for a review). This means that measuring events fix the positions of signals and these represent the positions of QM events. (Some physical particles such as photons are subject to little decoherence during propagation, even in water (cf: Anglin & Zurek (1996)).)

So signals in measuring devices usually have highly restrained positions. Now consider the final signals, the one's in phenomenal consciousness. To an observer of the brain they should be, very nearly, in their classical positions unless they consist of photons or are subject to some special effect such as has been proposed for microtubules. The brain acts as a measuring device causing decoherence. But despite this even signals composed of sodium ions, which should decohere rapidly in water, have a tiny, but finite, probability of remaining in a coherent state.

If your conscious experience is the signals and not the fabric of the brain are you the set of signals that interacts with the brain fabric almost immediately, the set that interacts after a minute or the set that almost never interacts? To an outside observer you must be the main chance, the rapidly interacting signals, but to the signals themselves all possibilities exist. Which one are you? Certainly any interaction between the signals and the mutually observed world must involve decoherence but the external observer would find it difficult to determine whether a particular interaction was due to signals that had interacted immediately or ones that were delayed (or delayed in an alternate QM reality). This problem is part of the *preferred basis problem* that will be discussed later.

Zurek (2003) assumes that phenomenal experience is identical to measurements. The observer is then both the signal and the apparatus that encloses the signal. He summarises the resultant idea of the completely determined observer who is fully integrated into the measured world:

The 'higher functions' of observers - e.g., consciousness, etc. - may be at present poorly understood, but it is safe to assume that they reflect physical processes in the information processing hardware of the brain. Hence, mental processes are in effect objective, as they leave an indelible imprint on the environment: The observer has no chance of perceiving either his memory, or any other macroscopic part of the Universe in some arbitrary superposition. " Zurek (2003)

Notice the phrase "perceiving .. his memory" - as neuroscientists we must ask "how"? By more measurements? There are no more measurements when things are arranged in phenomenal consciousness, the information has nowhere else to go. However, according to the empiricist philosophers the arrangements of the signals in phenomenal consciousness do extend through time in a definite order at any instant. Is it this order that determines the positions of signals in the brain or is it the brain that determines this order?

Physicalism leads us to an idea of the content of consciousness as an arrangement of quantum fields like the content of the brain or the content of the world. The arrangement of the quantum fields at an instant in experience is probably related to the arrangement of measured events at a succession of instants in the world.

The philosophy of qualia

The term "qualia" was introduced by C.I. Lewis in 1929:

This given element in a single experience of an object is what will be meant by "a presentation." Such a presentation is, obviously, an event and historically unique. But for most of the purposes of analyzing knowledge one presentation of a half-dollar held at right angles to the line of vision, etc., will be as good as another. If, then, I speak of " the presentation" of this or that, it will be on the supposition that the reader can provide his own illustration. No identification of the event itself with the repeatable content of it is intended.

In any presentation, this content is either a specific quale (such as the immediacy of redness or loudness) or something analyzable into a complex of such. The presentation as an event is, of course, unique, but the qualia which make it up are not. They are recognizable from one to another experience.(CI Lewis, *Mind and the World Order*, 1941 edition Chapter 2)

Tye (2003) gives the following definition of qualia:

"Experiences vary widely. For example, I run my fingers over sandpaper, smell a skunk, feel a sharp pain in my finger, seem to see bright purple, become extremely angry. In each of these cases, I am the subject of a mental state with a very distinctive subjective character. There is something it is like for me to undergo each state, some phenomenology that it has. Philosophers often use the term 'qualia' to refer to the introspectively accessible properties of experiences that characterize what it is like to have them. In this standard, broad sense of the term, it is very difficult to deny that there are qualia." Tye(2003).

In philosophy objects are considered to have perceived features such as shape and colour, weight and texture which are called sensible qualities. Sensible qualities are divided into

intrinsic, or primary, qualities that are properties of the object itself and extrinsic, or secondary, qualities which are related to the sensations produced in the observer. Shape is generally considered to be a primary quality whereas colour is often considered to be a secondary quality. It is generally considered that secondary qualities correspond to qualia (Smith 1990, Shoemaker 1990) and the two terms are often used synonymously. Although secondary qualities may be qualia, the term "qualia" may include things other than perceptions such as pain etc. that are, arguably, not secondary qualities. Primary qualities might also give rise to experience that is distinct from, say, the shape of an object itself.

Although "qualia" is a recent term, the philosophical debate about the nature of secondary qualities, such as colours, and the nature of conscious experience itself has been around for millenia.

It seems that the visual system gives rise to experience even in the absence of previous visual stimulation. For example, when someone recovers from blindness they have an experience that contains shapes and colours even though these have little meaning:

"When he first saw, he was so far from making any judgement of distances, that he thought all object whatever touched his eyes.... he knew not the shape of anything, nor any one thing from another, however different in shape and magnitude.. We thought he soon knew what pictures represented, which were shewed to him, but we found afterwards we were mistaken; for about two months after he was couched, he discovered at once they represented solid bodiess, when to that time he considered them only as party-coloured panes, or surfaces diversified with variety of paint." William Cheselden (1728)

Qualia are the components of experience, whatever the mode of input to that experience. Strawson (1994) includes content such as accompanies suddenly remembering or thinking of something as examples of qualia.

There is thought to be an *explanatory gap* associated with qualia (Levine 1983), as an example it is hard to imagine how the experience called pain could be a set of impulses in the brain.

Some philosophers have attempted to bridge this gap by invoking Direct Realism, proposing that our experience is in some way 'transparent' so that we experience the world or the injured limb directly (ie: there is an assumption that things flow within phenomenal experience into a centre point and we see right through this flow!). Strange though it may seem, this idea has led to a deduction that phenomenal experience is a set of things and qualities are these things, not deductions about or experiences of these things. As Tye (2003) puts it:

These observations suggest that qualia, conceived of as the immediately 'felt' qualities of experiences of which we are cognizant when we attend to them introspectively, do not really exist. The qualities of which we are aware are not qualities of experiences at all, but rather qualities that, if they are qualities of anything, are qualities of things in the world (as in the case of perceptual experiences) or of regions of our bodies (as in the case of bodily sensations). This is not to say that experiences do not have qualia. The point is that qualia are not qualities of experiences.

However, the outstanding issue for Tye's analysis is where in the world the thing that is called a quale exists - on a thing in the world beyond the body, on the retina, in the cortex, in the thalamus? Tye seems to be suggesting that "in the world" can only be beyond the

retina but given that a television can have a colour and a retina can have a colour why should we insist that the colour in conscious experience is always *of* the thing being represented via the DVD or videotape?

As was seen in the previous section, only signals are available in the classical world of conscious observation. The "reality" of the things that generate signals is not available. So whether experience is a signal at the position of what we call an "oak tree" or a signal in the eye due to photons reflected from the tree or a signal in the brain the same sort of phenomena would apply. Qualia would be a field of signals, not processes based on these signals.

Some philosophers hold that qualia are a field of signals derived from the original signals that are next to the quantum phenomena that compose an object. In other words they propose that qualia are not the first signals in the chain from whatever composes an object to the observer. These philosophers are known as Representationalists and the emphasis on secondary signals allows a contribution from the brain etc. to the field of signals that is conscious experience. Modern representationalists such as Tye (1995), Lehar(2003) and Dretske(2003) emphasise the idea that qualia are actual things that represent objects rather than concepts or experiences of things. As Dretske puts it:

"..the features that define what it is like to have an experience are properties that the objects we experience (not our experience of them) have.(Dretske 2003).

Lehar(2003) uses modern language to express the empiricist notion that the signals that comprise qualia are more likely to be in our brains than elsewhere, according to Lehar the objects we experience must be informational replicas in our heads:

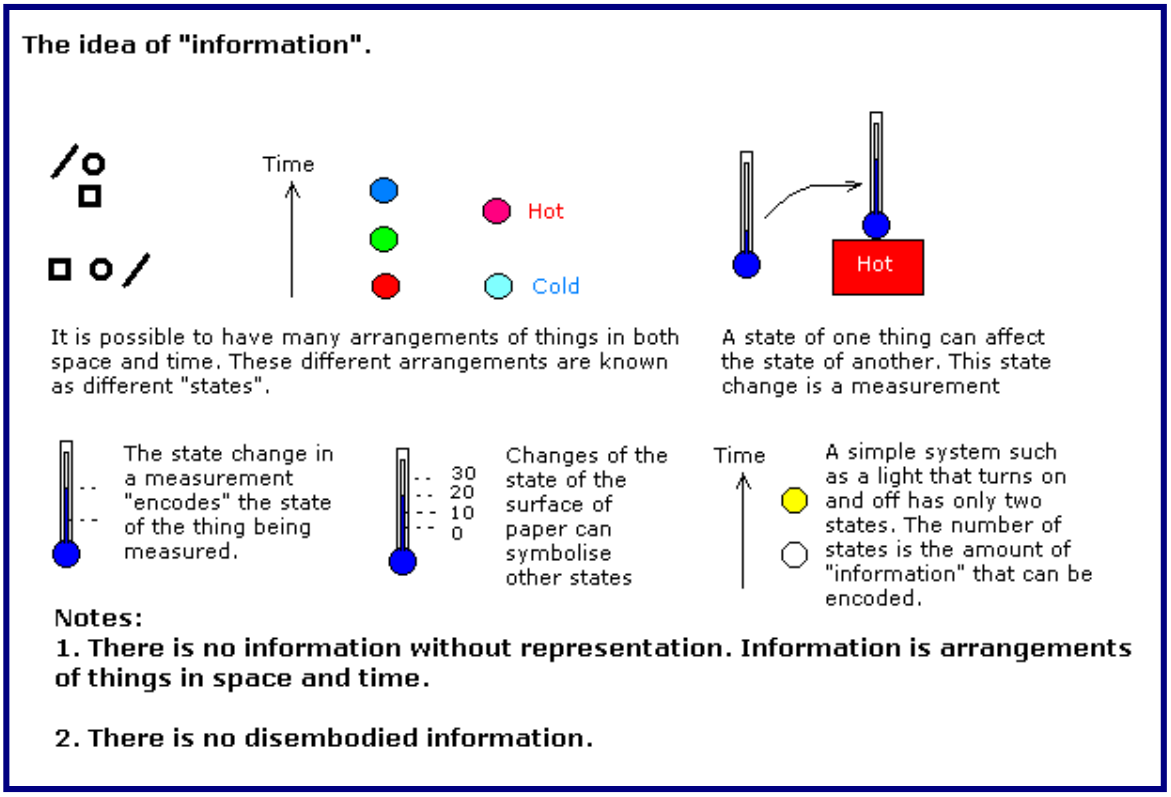
"The central message of Gestalt theory therefore is that the primary function of perceptual processing is the generation of a miniature, virtual-reality replica of the external world inside our head, and that the world we see around us is not the real external world, but is exactly that miniature internal replica (Lehar 2003)."

Direct Realists and Representationalists share the same view that qualia are an actual, physical field of things somewhere in the world. Some functionalists and eliminativists take a different view, believing that qualia do not exist except as judgements of properties that are used in interactions (ie: as disembodied information - see the section on Direct Realism).

Lewis, C.I. (1929) Mind and the World-Order. http://www.ditext.com/lewis/mwo2.html
Smith, A.D. (1990) Of Primary and Secondary Qualities, Philosophical Review 99 (1990).

Elementary Information and Information Systems Theory

When one physical thing interacts with another a change in "state" occurs. For instance, when a beam of white light, composed of a full spectrum of colours is reflected from a blue surface all colours except blue are absorbed and the light changes from white to blue. When this blue light interacts with an eye it causes blue sensitive cones to undergo a chemical change of state which causes the membrane of the cone to undergo an electrical change of state etc. The number of distinguishable states that a system can possess is the amount of information that can be encoded by the system.



Each distinguishable state is a "bit" of information. The binary symbols "1" and "0" have two states and can be used to encode two bits of information.

The binary system is useful because it is probably the simplest encoding of information and any object can represent a binary "1". In electrical digital systems an electrical pulse represents a "1" and the absence of a pulse represents a "0". Information can be transferred from place to place with these pulses. Things that transfer information from one place to another are known as "signals".

Symbols and information

0 1 2 3 4 5 6 7 8 9

The ten symbols of the decimal numbers are a symbol set that has ten states.

00 01 02 03 04 05 06 07 08 09
10 11 12 13 14 15 16 17 18 19
20 21 22 23 ..... etc.

Two occurrences of the symbols from the set can encode up to 100 states.

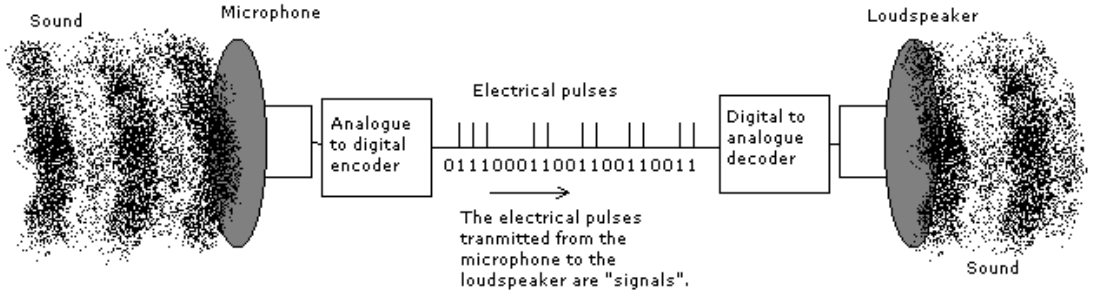
0 1 2 3
00 01 10 11

0 1 2 3 4 5 6 7
000 001 010 011 100 101 110 111

The simplest symbol set is the binary numbers. Two binary numbers can encode four states, three can encode up to eight states.

In science symbols are not disembodied. They are marks on paper or electrical impulses etc.

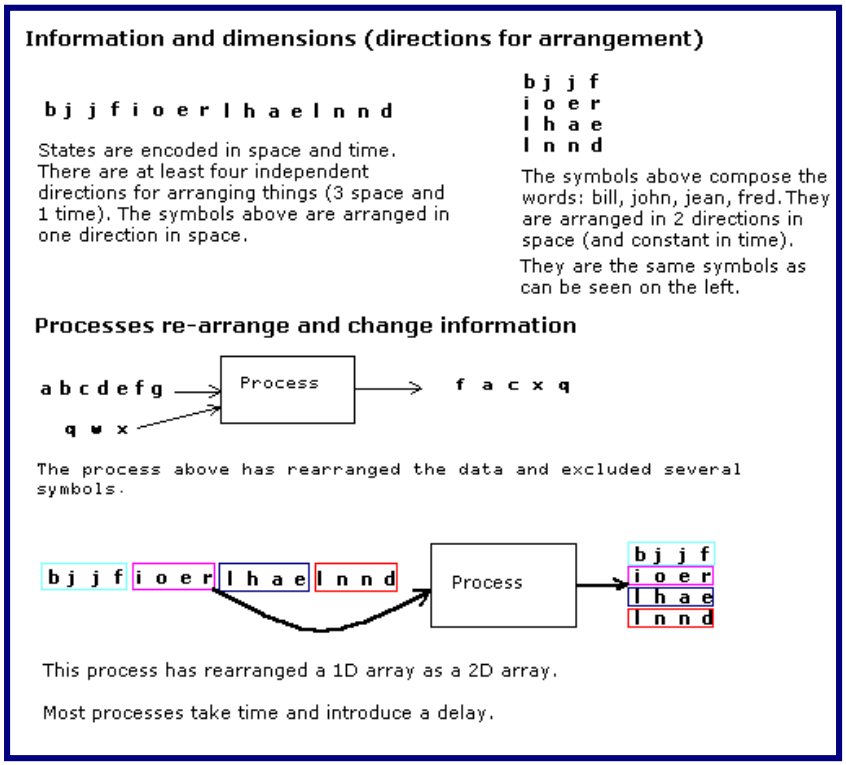
Signals and information



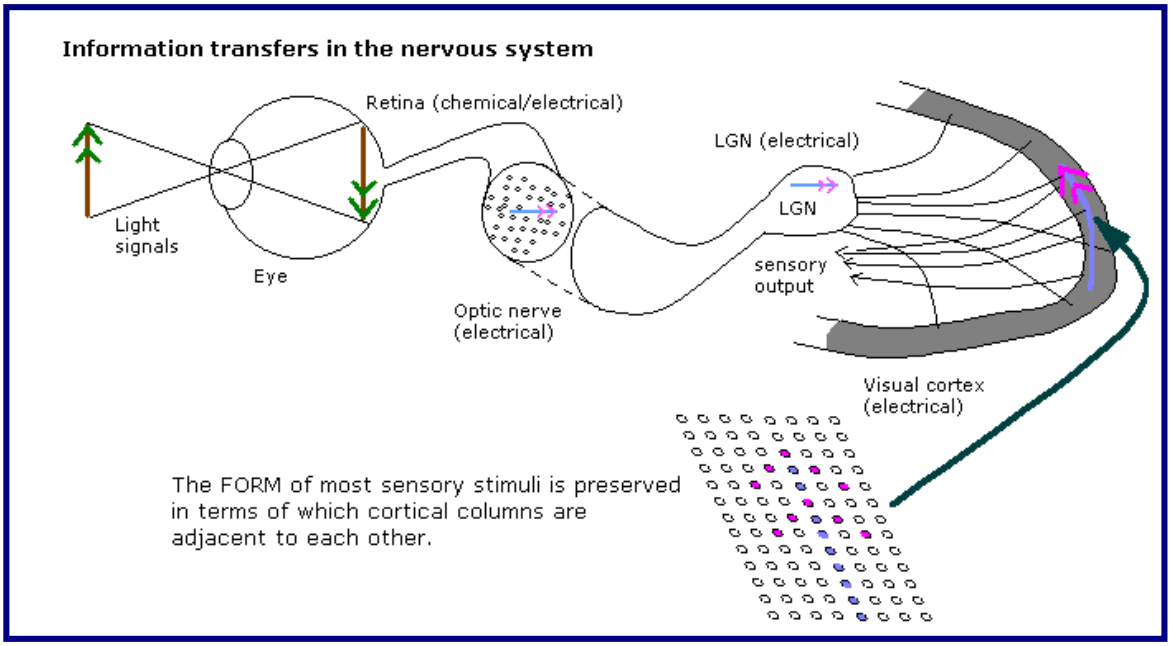
Information is encoded by changes of state, these changes can occur over time or as variations in density, temperature, colour etc. in the three directions in space. The writing on this page is spatially encoded.

It is interesting that our spoken communication uses a narrow band of sound waves. This favours the temporal encoding of information, in other words speech is largely a one dimensional stream of symbols. In vision, somesthesis, sound location and some of the other senses the brain uses spatial encoding of information as well as encoding over time.

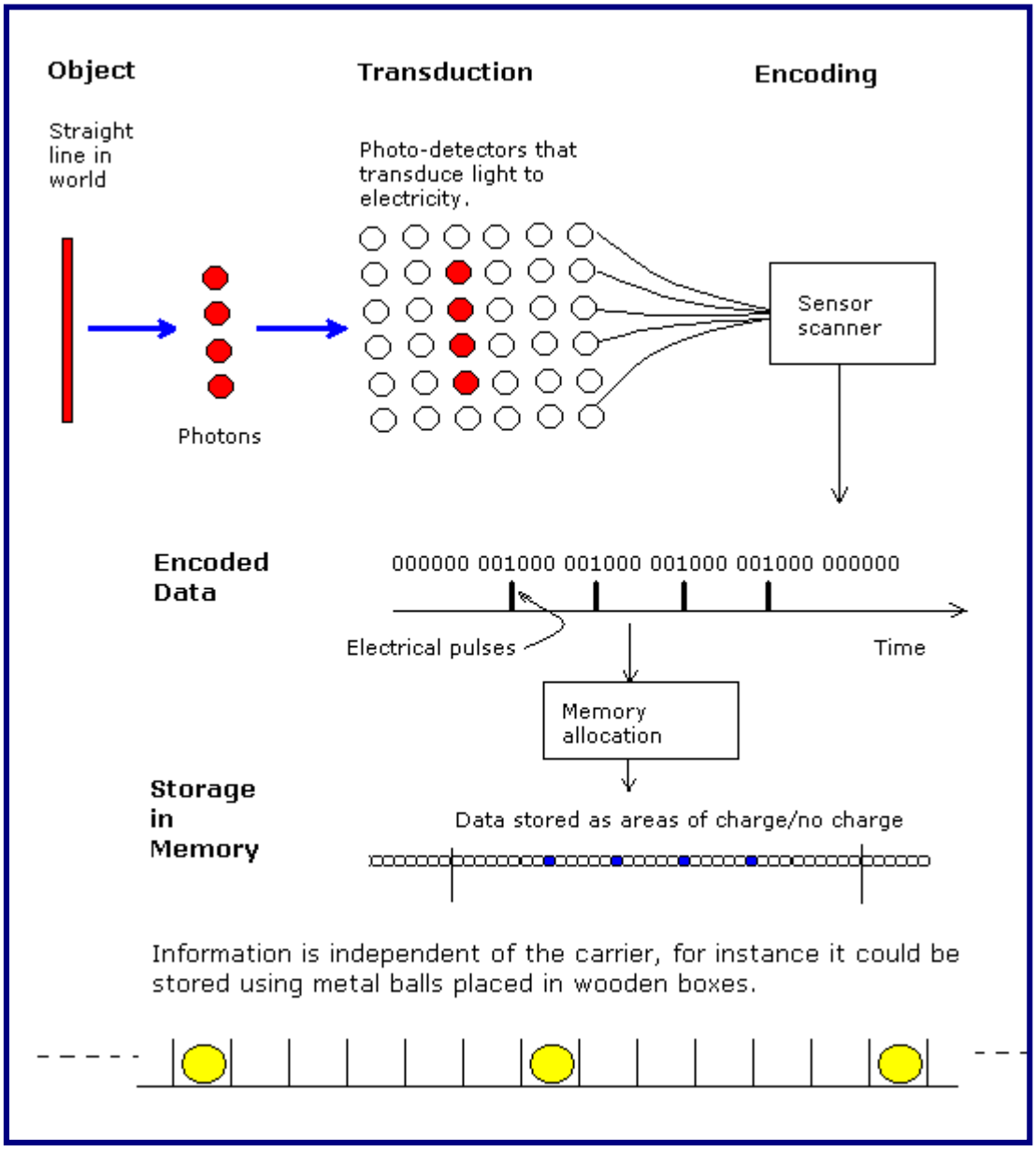
The illustration below shows how the directions available for arranging things can affect the encoding of information:



The rearrangement or replacement of a set of information so that some or all of the original information becomes encoded as another set of states is known as "processing". Devices that perform these actions are known as "information processors". The brain is predominantly an information processor.



Information systems in general have transducers that convert the state of signals in the world into signals impressed on another carrier, they then subject these signals to various processes and store them.



The spatial encoding in the brain generally preserves the relation of what is adjacent to what in the sensory field. This allows the form (geometry) of stimuli to be encoded.

Information transfers in the brain occur along numerous parallel "channels" and processes occur within each channel and between channels. Phenomenal consciousness at any moment contains a continuum of simultaneous (parallel) events. Classical processes take time so phenomenal experience is likely to be, at any instant, a simultaneous output of processes, not a classical process itself.

Classification, signs, sense, relations, supervenience etc.

A **sign** is a symbol, combination of symbols such as a **word** or a combination of words. A **referent** is "...that to which the sign refers, which may be called the reference of the sign" (Frege 1892). Statements and concepts usually express relations between referents.

The **sense** of statements depends on more than the simple referents within them, for instance "the morning star is the evening star" is true in terms of the referents but dubious in terms of the sense of the morning and evening stars because the morning star is Venus as seen in the morning and the evening star is Venus as seen in the evening. So the sense of the expression "the morning star" depends on both the referent "Venus" and the referent "Morning" and probably other associations such as "sunrise", "mist" etc..

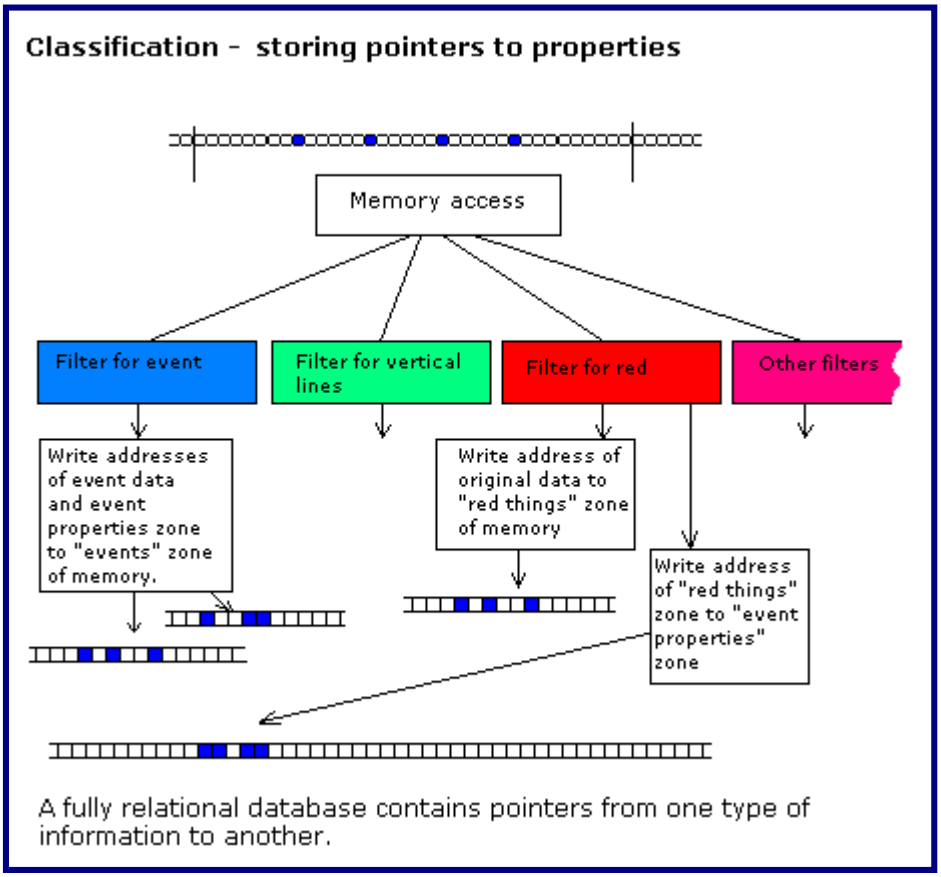
Each sign is related to many other signs and it is these groups of relationships that provide the sense of a sign or a set of signs. A **relation** is an association between things. It can be understood in the abstract as "what is next to what". Relations occur in both time and space. When a ball bounces the impact with the floor changes the direction of the ball so "direction" is related to "impact", the ball is round so "ball" is related to "round". For instance, the morning is next to the presence of the morning star so "morning" and "morning star" are related. Relations are the connections that allow classification.

According to the physical concept of information all abstract signs are physical states of a signal and are only abstract according to whether they are related to a physical thing or exclusively to another sign. The process of treating an abstract idea as if it were a concrete thing that contains other concrete things is known as **reification**.

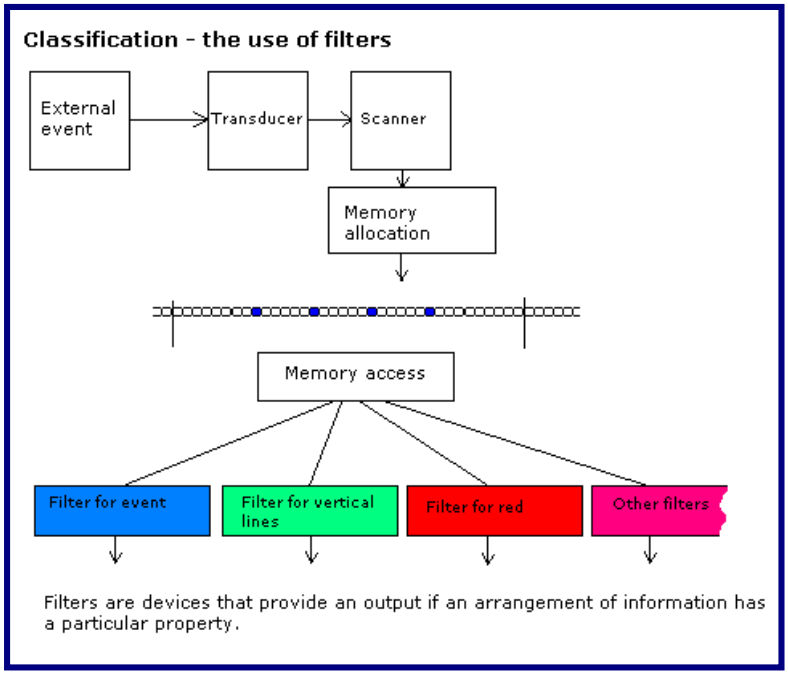
It is possible to have statements that have a sense but apparently no reference. As Frege put it, the words 'the celestial body most distant from the Earth' have a sense but may not have a reference. There can be classes of things that have not yet acquired any members or have no members. In a physical sense a particular **class** is a sign that refers to a particular state or set of states. Classes can be arbitrary such as "big things" being all things that have a state of being over one metre long. Classes and **sets** are very similar, sometimes sets are defined as being a class that is an element of another class. The term "set" has largely superceded the term "class" in academic publications since the mid twentieth century.

The **intension** of a set is its description or defining properties. The **extension** of a set is its members or contents. In mathematics a set is simply its members, or extension. In philosophy there is considerable discussion of the way that a given description can describe more than one thing. In other words, one intension can have several extensions. The set of things that are "tables" has the properties "legs", "flat surface" etc. The extension of "tables" is all the physical tables. The intension of "tables" may also include "stools" unless there is further clarification of the properties of "tables". Intensions are functions that identify the extensions (original members of a set) from the properties.

Classification is performed by information systems and by the information processing parts



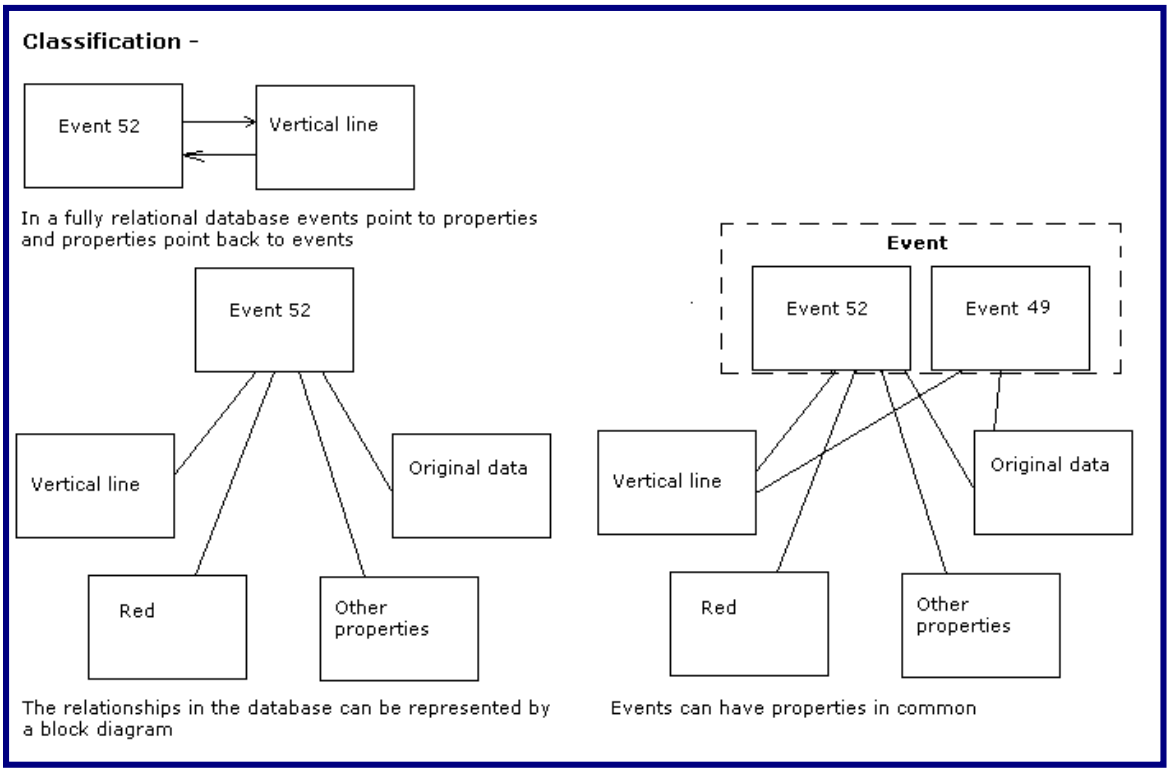
of the nervous system. A simple classification is to sort symbols according to a set of rules, for instance a simple **sort** classifies words by letter sequence. There are numerous classification systems in the visual system such as arrangements of neurons that produce a single output when a particular orientation of a line is viewed or a particular face is seen etc.



The processes that identify attributes and properties of a thing are usually called **filters**. The output of filters becomes the properties of a set and specifies the relations between sets. These relations are stored as address pointers in computers or connections in the nervous system.

An intension uses these properties and relations to identify the things that are members of the set in the world. Clearly the more specific the filters the more accurate the intension.

A **database** is a collection of signs. A **fully relational** database is a database arranged in related sets with all relationships represented by pointers or connections. In conventional usage a **relational database** is similar but more sophisticated, redundant relationships and wasteful storage being avoided. Conventional relational databases obey "Codd's laws". An **hierarchical database** only contains pointers that point from the top of a classification hierarchy downwards. Events and persistent objects are also known as **entities**, the output of filters related to entities are known as the **attributes** of the entity. In practice a system requires an event filter to record an entity (in a computer system the event filter is usually a single data entry form and the attributes are filtered using boxes on the screen to receive typed input).



In information systems design there are many ways of representing **classification hierarchies**, the most common is the **entity diagram** which assumes that the attributes of an entity define it and are stored together physically with the symbols that represent the entity. This adjacent storage is purely for convenient management of storage space and reduction of the time required for retrieval in modern computers.

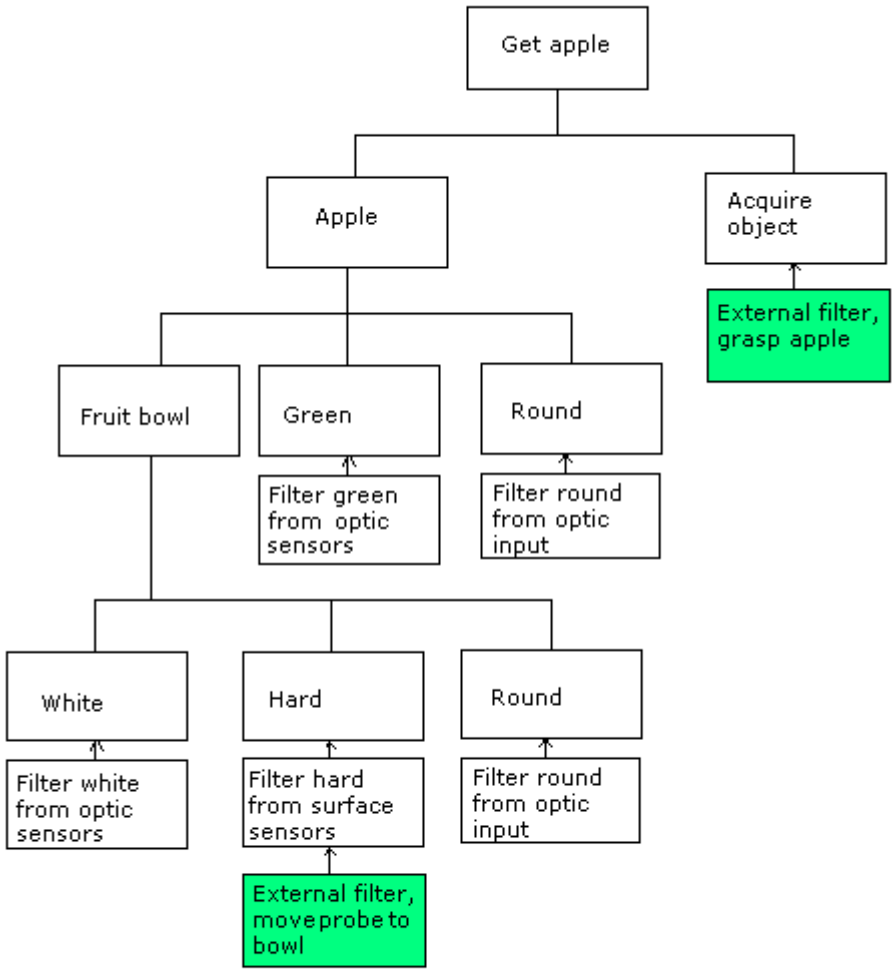
Filters contain processing agents of varying degrees of sophistication from simple sorting processes to "intelligent" processes such as programs and neural networks. It is also

possible to arrange filters in the world beyond an information processor. For instance, an automatic text reading machine might turn over the pages of a book to acquire a particular page. A human being might stroke an object to confirm that the texture is as it appears to be and so on.

Scientists routinely use external transducers and filters for the purpose of classification. For instance, a mass spectrometer could be used to supply details of the atomic composition of an item. External filters allow us to distinguish between things that are otherwise identical (such as two watery compounds XYZ and H2O) or to acquire properties that are unobservable with biological transducers such as the eyes and ears. The scientist plus his instruments is a single information system. In practice the referent of a set is determined by applying transducers and filters to the world and looking up the results in a relational database. If the result is the original set then a referent has been found. A sophisticated system may apply "fuzzy logic" or other methods to assign a probability that an object is truly a member of a particular set.

It is also possible to classify information according to relationships in time (ie: starting a car's engine is related to car moving away). Within an information system the output from the filter for "starting engine" might precede that from the filter for "starts moving". In information systems design procedures that involve successions of events can be arranged in classification structures in the same way as data; this technique is known as **structured programming** (esp. Jackson structured programming).

Structured information processing.



A highly simplified and schematic block diagram of information processing where data and procedures are combined in a classification structure. Processing proceeds from bottom left to the right and upwards.

Hierarchies related to a single entity are frequently stored together as **objects** and the information processing that results is known as **object oriented programming**. A fully relational database would, in principle, contain all the objects used in a structured information system.

It has been pointed out by (McCarthy and Hayes (1969)) that an information processor that interacts with the environment will be producing continuous changes in all of its classifications (such as position etc) and also changes in theories (structured programs that are predictive processes) about the world. In a serial processor, such as a Turing Machine with a one dimensional tape, the presence of changes in the world would create a huge burden on the machine. In a parallel processor, such as a biological neural network, the reclassifications should be straightforward. The problem of adapting an information system to changes in the world, most of which have little effect on the processes performed by the system, is known as the **frame problem**. The frame problem is usually stated in a form such as "how is it possible to write formulae that describe the effects of actions without having to write a large number of accompanying formulae that describe the mundane, obvious non-effects of those actions?" (Shanahan 2004).

Chalmers(1996) introduced the terms **primary intension** and **secondary intension**. Primary intension is a high level description where the properties of a set may be insufficient to specify the contents of the set in the physical world. For instance, the term "watery" might specify several liquids with various compositions. Secondary intension is specific so that it applies to one substance in the world (H2O). In the context of information systems primary intensions differ from secondary intensions as a result of inadequate filtering and classification. (See note below for details of Putnam's twin earth thought experiement).

The problem of matching the properties and relations of an item in a relational database with an item in the world involves the problem of **supervenience**. Supervenience occurs when the properties and relations in the database for an item are the same as the output from filters applied to the item. In other words, **in an information system information does not supervene directly on an object, it supervenes on information derived from the object**. Chalmers described supervenience in terms that are accessible to an information systems approach:

"The properties of A supervene on the properties of B if no two possible situations are identical with respect to the properties of A while differing with respect to the properties of B (after Chalmers 1996)."

In terms of information processing the *properties* are changes in state derived from a transducer that are subject to classification with a filter. The properties of a predictive program would supervene on the input from transducers applied to an object if it correctly identified the sets and sequence of sets that are discovered at all times.

The reader should be cautioned that there is an extensive literature associated with supervenience that does not stress the way that information is embodied and representational. (The removal of these constraints will lead to non-physical theories of information).

It is sometimes asked how conscious experience containing a quale that is a colour, such as

blueness, can supervene on the physical world. In terms of information systems the question is back to front: blueness is very probably a phenomenon in the physical brain - it is certainly unlike an arrangement of stored bits in an information system. The question should read "what physical theory supervenes on information in the signals related to the phenomenon called blue?"

The simple answer is that there is no widely accepted description available of the physical nature of the experience called blue (there are several theories however). A common mistake is to say that the secondary intension of the quale blue is known - this is not the case, the physical basis of em radiation or absorption of light is known to some extent but these are almost certainly not the physical basis of the "blue" of experience. The quale "blue" is probably a particular substrate that has a state, not an encoded state on a generalised substrate.

Information is the patterns and states of an underlying substrate or carrier, this leaves us with exciting questions such as: what is it like to be the substrate itself rather than simply the information impressed upon it? Can only particular substrates constitute conscious experience? How can we relate the properties of this experience to information about the physical world?

The substrate of information is not part of the problem of access consciousness that deals with the problem of the flow of information from place to place.

- Frege, G. (1892) On Sense and Reference. http://en.wikisource.org/wiki/On_Sense_and_Reference

- Pruss, A.R. (2001) The Actual and the Possible. in Richard M. Gale (ed.), Blackwell Guide to Metaphysics, Oxford: Blackwell. http://www.georgetown.edu/faculty/ap85/papers/ActualAndPossible.html

- Menzies, P (2001). Counterfactual Theories of Causation. Stanford Encyclopedia of Philosophy http://plato.stanford.edu/entries/causation-counterfactual/

- McCarthy, J. and Hayes, P.J. (1969), "Some Philosophical Problems from the Standpoint of Artificial Intelligence", Machine Intelligence 4, ed. D.Michie and B.Meltzer, Edinburgh: Edinburgh University Press, pp. 463-502.

- Shanahan, M. (2004) "The frame problem". Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/frame-problem/

Qualia and Information

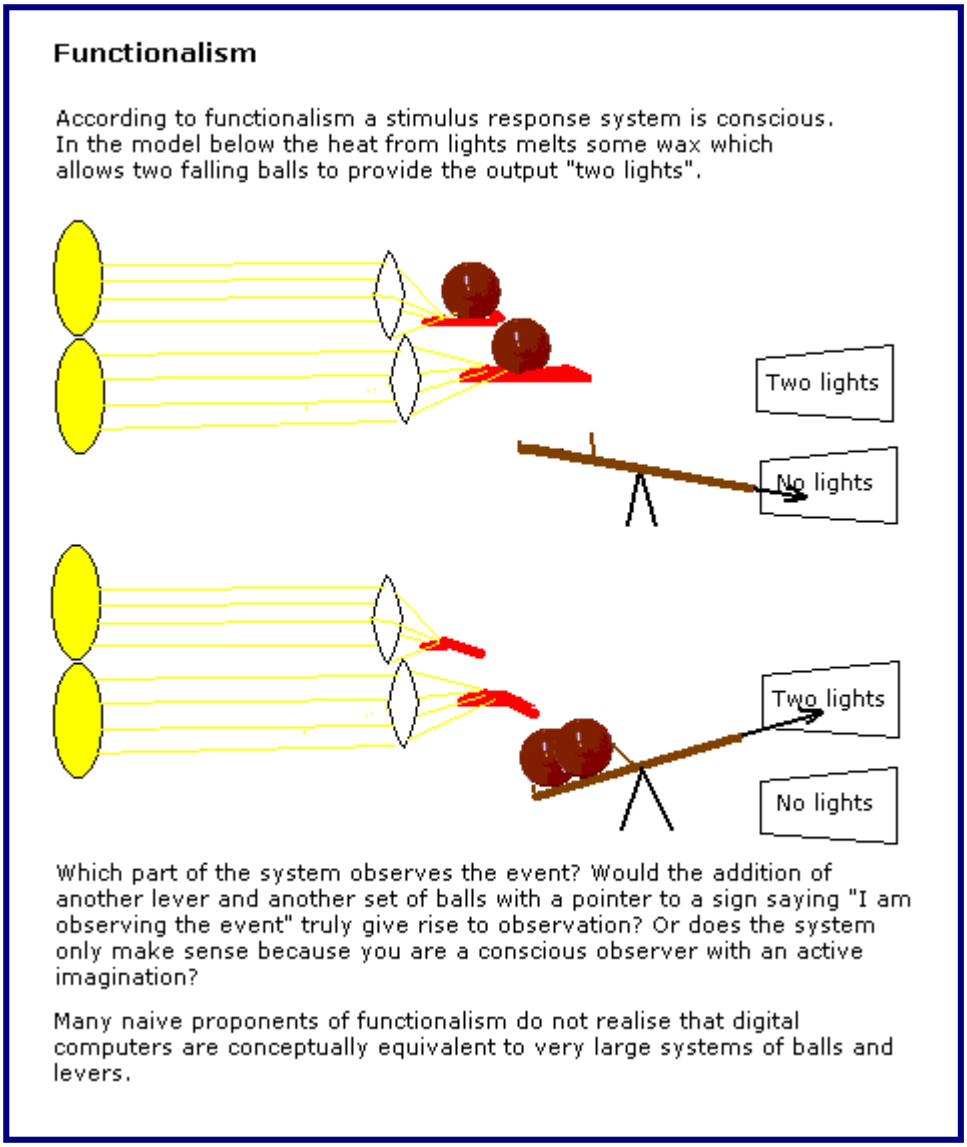
The problem of the generalised nature of information is addressed by several "thought experiments" which are described below.

The problem of "intensions" is tackled in Putnam's twin earth thought experiment which was discussed above but is given in more detail below.

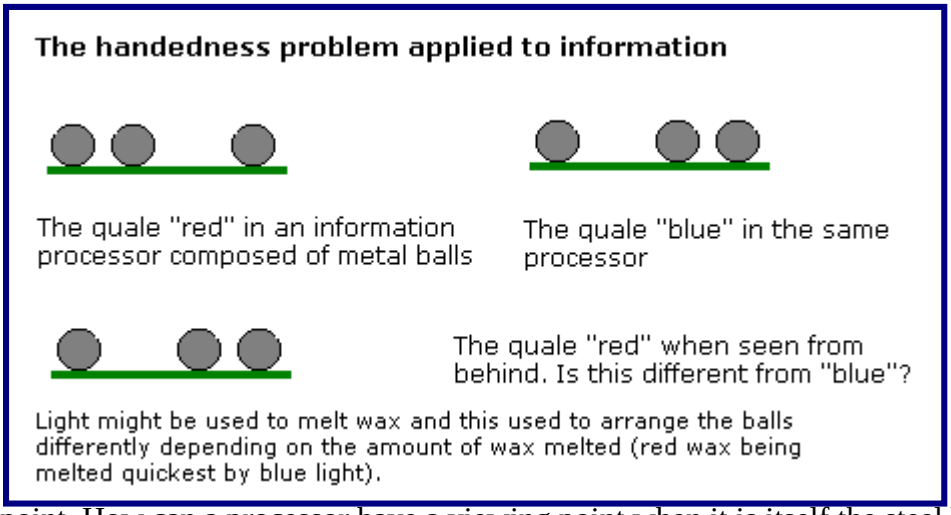
Absent and fading qualia

Absent qualia

Block (1978) argued that the same functions can be performed by a wide range of systems. For instance, if the population of China were equipped with communication devices and a set of rules they could perform almost any function but would they have qualia? The argument considers the fact that systems which process information can be constructed of a wide range of materials and asks whether such systems will also have qualia (see illustration below).



This argument also occurs when the physical structure of computing devices is considered, for instance a computing machine could be constructed from rolling steel balls. Would the steel balls at one instant possess the quale 'blue' and then, as a result of the movement of one ball to another position, possess the quale 'red'? Can an arrangement of balls really have qualia or are they absent? It is encumbent upon proponents of functional organisation to describe why identical balls arranged as O O OOO can be the quale red and yet those arranged as OOO O O can be the quale blue. They must also take into account Kant's "handedness problem": the balls OOO O O look like O O OOO when viewed from behind. Red and blue, as arrangements of things, would be identical depending on the viewing



point. How can a processor have a viewing point when it is itself the steel balls?

Fading qualia

Pylyshyn (1980) introduced a thought experiment in which a human brain is progressively replaced by synthetic components and it is asked what would happen to consciousness during this replacement of the brain.

Chalmers (1996) considers the problem in depth from the point of view of functional organisation. (ie: considering replacement of biological components with components that perform the same functions). The argument is straightforward: if phenomenal consciousness is due to functional organisation then replacement of biological parts with artificial parts that duplicate the function should allow phenomenal consciousness to continue.

But suppose phenomenal consciousness is not due to functional organisation. What would we expect then?

Chalmers argues that consciousness could not suddenly disappear during replacement of the brain because functions could be replaced in tiny stages so unless qualia could reside in a single tiny place in the brain **Disappearing qualia** would be ruled out. This denial that qualia could be in a small physical place is equivalent to denying that qualia could be a continuous field of things and is highly contentious. Those who hold that qualia are continuous might maintain that progressive replacement of the brain could indeed result in one neuron that has phenomenal qualia in it or around it in the form of a field.

If qualia are due to something other than functions then there would be a very real possibility that the resolution of conscious experience would decrease as the data available decreases. This possibility was rejected by Chalmers because it implies that conscious experience could occur in a very small field.

Chalmers considers the alternative idea of **fading qualia** where slow replacement of parts reduces experience progressively. This "fading" is described in terms of qualia fading from red to pink and experience in general becoming more and more out of step with the world. Chalmers summarily dismisses the idea of fading qualia on the grounds that people do not have abnormal experiences, like fading colours, except in the case of pathology. This is an odd argument considering that few things could be as pathological as replacing the brain with inorganic parts.

It is possible that at some stage during the replacement process the synthetic parts alone would have sufficient data to identify objects and properties of objects so that the experience would be like blindsight. The subject might be amazed that subjective vision was disappearing. However, Chalmers denies that new beliefs, such as amazement at a new state, would be possible. He says that:

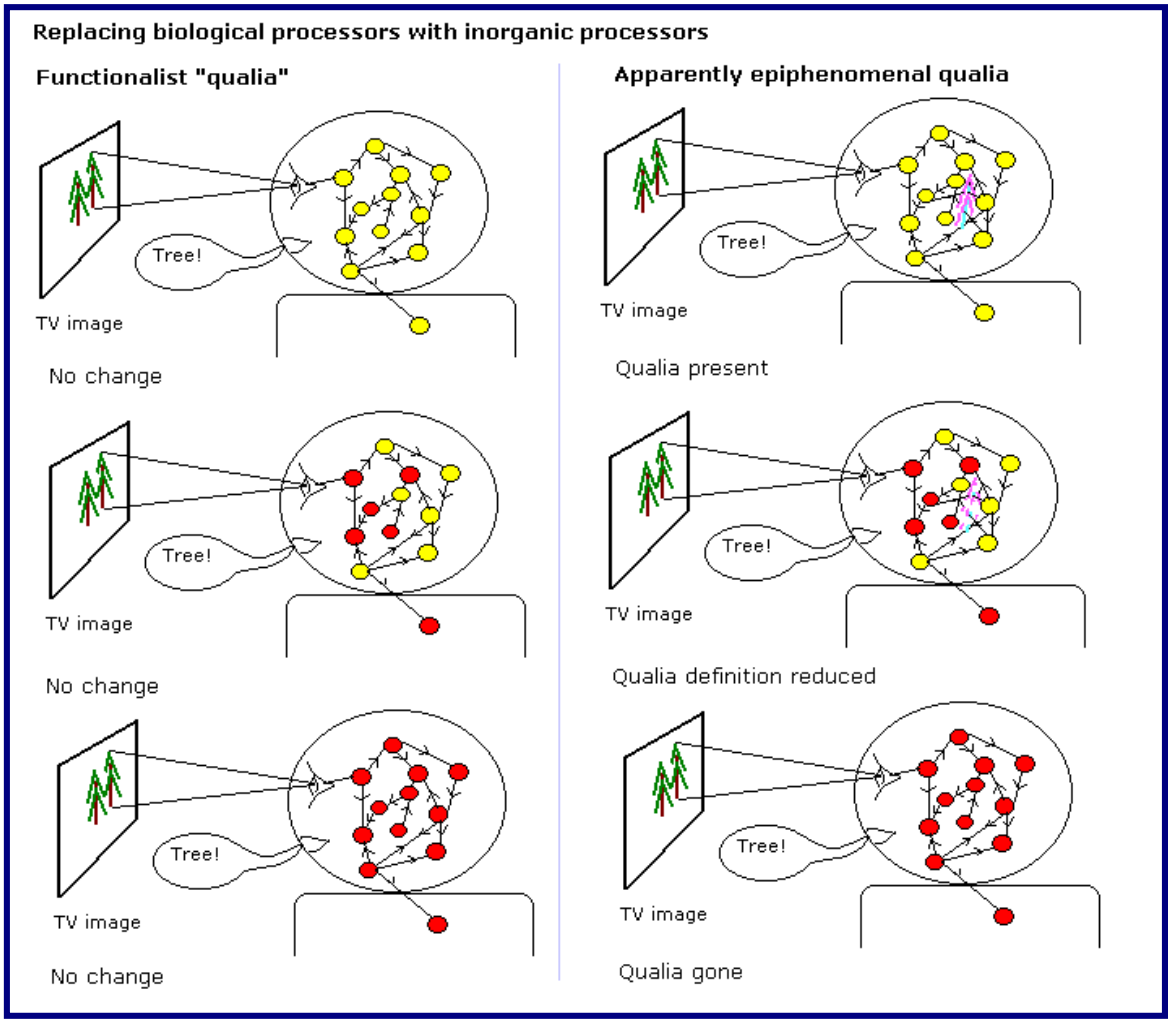
"Nothing in the physical system can correspond to that amazement. There is no room for new beliefs such as "I can't see anything," new desires such as the desire to cry out, and other new cognitive states such as amazement. Nothing in the physical system can correspond to that amazement."

On the basis of the impossibility of new beliefs Chalmers concludes that fading qualia are impossible. This is strange because there is no reason why a non-conscious information processor attached to a conscious area of brain should not report that certain types of data are absent from it. The reporting of "I can't see anything" (ie: "I have no conscious experience") would show that there was no conscious part of the brain remaining.

According to Chalmers, if fading qualia do not occur then qualia must also exist in "Robot", a totally synthetic entity, so **absent qualia** do not occur either. Therefore Robot should be conscious. He concludes the fading qualia argument by stating that it supports his theory that consciousness results from *organizational invariance*, a specific set of functions organised in a particular way:

"The invariance principle taken alone is compatible with the solipsistic thesis that my organization gives rise to experience. But one can imagine a gradual change to my organization, just as we imagined a gradual change to my physical makeup, under which my beliefs about my experience would be mostly preserved throughout, I would remain a rational system, and so on. For similar reasons to the above, it seems very likely that conscious experience would be preserved in such a transition"

Chalmers is aware that if qualia were epiphenomenal the argument might be false.


Replacing biological processors with inorganic processors
Functionalist "qualia"
Apparently epiphenomenal qualia
Tree!
TV image
No change
Qualia present
Qualia definition reduced
Qualia gone

There are other interesting questions related to the fading qualia argument, for instance: Can all of organic chemistry be replaced by inorganic chemistry - if not why not? If information always has a physical substrate and conscious experience is the arrangement of that substrate then how could conscious experience be the same if the substrate is replaced? At the level of molecular and atomic interactions almost all functions involve electromagnetic fields, if identical function is achieved at scales below the size of an organelle in a cell in the brain would the functional elements, such as electromagnetic fields, have been changed?

The reader may have spotted that Chalmers' fading qualia argument is very similar to Dennett's argument about the non-existence of qualia. In Dennett's argument qualia are dubiously identified with judgements and then said to be non-existent. In Chalmer's argument an attempt is made to identify qualia with beliefs about qualia so they can be encompassed by a functionalist theory.

Pylyshyn, Z. (1980) The "causal power" of machines. Behavioral and Brain Sciences 3:442-444.

Chalmers, D.J. (1996). The Conscious Mind. Oxford University Press.

Putnam's twin earth thought experiment

The original Twin Earth thought experiment was presented by philosopher Hilary Putnam in his important 1975 paper "The Meaning of 'Meaning'", as an early argument for what has subsequently come to known as semantic externalism. Since that time, philosophers have proposed a number of variations on this particular thought experiment, which can be collectively referred to as Twin Earth thought experiments.

Putnam's original formulation of the experiment was this:

We begin by supposing that elsewhere in the universe there is a planet exactly like earth in virtually all respects, which we refer to as 'Twin Earth'. (We should also suppose that the relevant surroundings of Twin Earth are identical to those of earth; it revolves around a star that appears to be exactly like our sun, and so on.) On Twin Earth there is a Twin equivalent of every person and thing here on Earth. The one difference between the two planets is that there is no water on Twin Earth. In its place there is a liquid that is superficially identical, but is chemically different, being composed not of H2O, but rather of some more complicated formula which we abbreviate as 'XYZ'. The Twin Earthlings who refer to their language as 'English' call XYZ 'water'. Finally, we set the date of our thought experiment to be several centuries ago, when the residents of Earth and Twin Earth would have no means of knowing that the liquids they called 'water' were H2O and XYZ respectively. The experience of people on Earth with water, and that of those on Twin Earth with XYZ would be identical.

Now the question arises: when an earthling, say Oscar, and his twin on Twin Earth (also called 'Oscar' on his own planet, of course. Indeed, the inhabitants of that planet necessarily call their own planet 'earth'. For convenience, we refer to this putative planet as 'Twin Earth', and extend this naming convention to the objects and people that inhabit it, in this case referring to Oscar's twin as Twin-Oscar, or Toscar.) say 'water' do they mean the same thing? Ex hypothesi, their brains are molecule-for-molecule identical. Yet, at least according to Putnam, when Oscar says water, the term refers to H2O, whereas when Toscar says 'water' it refers to XYZ. The result of this is that the contents of a persons brain are not

sufficient to determine the reference of terms he uses, as one must also examine the causal history that led to his acquiring the term. (Oscar, for instance, learned the word 'water' in a world filled with H2O, whereas Toscar learned 'water' in a world filled with XYZ.) This is the essential thesis of semantic externalism. Putnam famously summarized this conclusion with the statement that "meaning just ain't in the head."

In terms of physical information systems such as occur in the brain this philosophical argument means that if there are inadequate external filters available the information system will confuse XYZ with H2O; it will conclude that they are they same thing and have no difference in meaning. For the information system meaning is in the classification structures assigned by the system. If the system is provided with better transducers and filters then new meanings will arise within the system. However, for an information system 'meaning' is no more than a chain of relations because this is the nature of information (ie: arrangements of an arbitrary carrier). Other types of meaning would require phenomena other than simple information processing.

In Putnam's thought experiment the world can be different but the meaning for the individual is the same if the brain is the same. If there is a type of meaning other than a chain of relations would Putnam's experiment suggest that this type of 'meaning' occurs as a phenomenon in the brain or in the world beyond the body?

Putnam, H. (1975/1985) The meaning of 'meaning'. In Philosophical Papers, Vol. 2: Mind, Language and Reality. Cambridge University Press.

The Inverted Qualia Argument

The possibility that we may each experience different colours when confronted by a visual stimulus is well known and was discussed by John Locke. In particular the idea of **spectrum inversion** in which the spectrum is exchanged, blue for red and so on is often considered. It is then asked whether the subject of such an exchange would notice any difference. Unfortunately it turns out that colour is not solely due to the spectrum and depends on hue, saturation and lightness. If the colours are inverted all the axes of colour would need to be exchanged and the relations between the colours would indeed still be discernably different.

Some philosophers have tried to avoid this difficulty by asking questions about qualia when the subject has no colour vision. For instance, it is asked whether a subject who saw things in black and white would see the world differently from one who saw the world in white and black.

This sort of discussion has been used as an attack on Behaviourism where it is argued that whether a tomato is seen as black or white the subject's behaviour towards the tomato will be the same. So subject's can have mental states independent of behaviours.

Block (1990) has adapted this argument to an *inverted earth* scenario in which it is proposed that a subject goes to another planet which is identical to earth except for the inversion of visual qualia. He points out that behaviours would adjust to be the same on the inverted earth as on the actual earth. All functions would be identical but the mental state would be different so it is concluded that mental states are not processes.

Chalmers(1996) approaches this argument by assuming that the absent and fading qualia

arguments have proven his idea of organisational invariance. He then introduces the idea that conscious experience only exists for the durationless instant and notes that, given these assumptions a person would not be aware that the quale red had been switched for the quale blue.

"My experiences are switching from red to blue, but *I do not notice any change*. Even as we flip the switch a number of times and my qualia dance back and forth, I will simply go about my business, noticing nothing unusual."

Block, N. (1990). Inverted Earth, Philosophical Perspectives, 4: 53-79.

See also: Block, N. Qualia. http://www.nyu.edu/gsas/dept/philo/faculty/block/papers/qualiagregory.pdf Byrne, A. (2004). Inverted Qualia. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/qualia-inverted/ Shoemaker, S. (2002). CONTENT, CHARACTER, AND COLOR II: A BETTER KIND OF REPRESENTATIONALISM Second Whitehead Lecture. http://humanities.ucsc.edu/NEH/shoemaker2.htm

The Knowledge Argument

Much of the philosophical literature about qualia has revolved around the debate between physicalism and non-physicalism. In 1982 Frank Jackson proposed the famous "Knowledge Argument" to highlight how physical knowledge might not be enough to describe phenomenal experience:

"Mary is a brilliant scientist who is, for whatever reason, forced to investigate the world from a black and white room via a black and white television monitor. She specializes in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like 'red', 'blue', and so on. She discovers, for example, just which wavelength combinations from the sky stimulate the retina, and exactly how this produces via the central nervous system the contraction of the vocal chords and expulsion of air from the lungs that results in the uttering of the sentence 'The sky is blue'. (It can hardly be denied that it is in principle possible to obtain all this physical information from black and white television, otherwise the Open University would of necessity need to use color television.)

What will happen when Mary is released from her black and white room or is given a color television monitor? Will she learn anything or not? It seems just obvious that she will learn something about the world and our visual experience of it. But then it is inescapable that her previous knowledge was incomplete. But she had all the physical information. Ergo there is more to have than that, and Physicalism is false. Jackson (1982).

The Knowledge argument is a category mistake because a description of the universe, such as information about science, is a set of symbols in a particular medium such as ink on paper. These symbols provide the recipe for experiments and other manipulations of nature, and predict the outcome of these manipulations. The manipulations of nature are not the same as the set of symbols describing how to perform these manipulations. Scientific information is not the world itself and the truth or falsehood of Physicalism is unaffected by the knowledge argument.

If the Knowledge Argument is interpreted as an argument about whether information about the nature of the colour red could ever be sufficient to provide the experience that we call

red then it becomes more relevant to the problem of consciousness but it is then a debate about whether information processors could be conscious, this is covered below. Those interested in a full discussion of the Knowledge Argument should consult Alter (1998) and especially the link given with this reference.

The problem of machine and digital consciousness

Information processing and digital computers

Information processing consists of encoding a state, such as the geometry of an image, on a carrier such as a stream of electrons, and then submitting this encoded state to a series of transformations specified by a set of instructions called a program. In principle the carrier could be anything, even steel balls or onions, and the machine that implements the instructions need not be electronic, it could be mechanical or fluidic.

Digital computers implement information processing. From the earliest days of digital computers people have suggested that these devices may one day be conscious. One of the earliest workers to consider this idea seriously was Alan Turing. Turing proposed the Turing Test as a way of discovering whether a machine can think. In the Turing Test a group of people would ask a machine questions and if they could not tell the difference between the replies of the machine and the replies of a person it would be concluded that the machine could indeed think. Turing's proposal is often confused with the idea of a test for consciousness. However, phenomenal consciousness is an internal state so the best that such a test could demonstrate is that a digital computer could simulate consciousness.

If technologists were limited to the use of the principles of digital computing when creating a conscious entity they would have the problems associated with the philosophy of 'strong' artificial intelligence. The term *strong AI* was defined by Searle:

..according to strong AI, the computer is not merely a tool in the study of the mind; rather, the appropriately programmed computer really is a mind (J. Searle in Minds, Brains and Programs. The Behavioral and Brain Sciences, vol. 3, 1980).

If a computer could demonstrate Strong AI it would not necessarily be more powerful at calculating or solving problems than a computer that demonstrated Weak AI.

The most serious problem with Strong AI is John Searle's "chinese room argument" in which it is demonstrated that the contents of an information processor have no intrinsic meaning -at any moment they are just a set of electrons or steel balls etc. The argument is reproduced in full below:

"One way to test any theory of the mind is to ask oneself what it would be like if my mind actually worked on the principles that the theory says all minds work on. Let us apply this test to the Schank program with the following Gedankenexperiment. Suppose that I'm locked in a room and given a large batch of Chinese writing. Suppose furthermore (as is indeed the case) that I know no Chinese, either written or spoken, and that I'm not even confident that I could recognize Chinese writing as Chinese writing distinct from, say, Japanese writing or meaningless squiggles. To me, Chinese writing is just so many meaningless squiggles. Now suppose further that after this first batch of Chinese writing I am given a second batch of Chinese script together with a set of rules for correlating the second batch with the first batch. The rules are in English, and I understand these rules as well as any other native speaker of English. They enable me to correlate one set of formal symbols with another set of formal symbols, and all that "formal" means here is that I can identify the symbols entirely by their shapes. Now suppose also that I am given a third batch of Chinese symbols together with some instructions, again in English, that enable me to correlate elements of this third batch with the first two batches, and these rules instruct

me how to give back certain Chinese symbols with certain sorts of shapes in response to certain sorts of shapes given me in the third batch. Unknown to me, the people who are giving me all of these symbols call the first batch a "script," they call the second batch a "story," and they call the third batch "questions." Furthermore, they call the symbols I give them back in response to the third batch "answers to the questions," and the set of rules in English that they gave me, they call the "program." Now just to complicate the story a little, imagine that these people also give me stories in English, which I understand, and they then ask me questions in English about these stories, and I give them back answers in English. Suppose also that after a while I get so good at following the instructions for manipulating the Chinese symbols and the programmers get so good at writing the programs that from the external point of view "that is, from tile point of view of somebody outside the room in which I am locked" my answers to the questions are absolutely indistinguishable from those of native Chinese speakers. Nobody just looking at my answers can tell that I don't speak a word of Chinese. Let us also suppose that my answers to the English questions are, as they no doubt would be, indistinguishable from those of other native English speakers, for the simple reason that I am a native English speaker. From the external point of view "from the point of view of someone reading my answers" the answers to the Chinese questions and the English questions are equally good. But in the Chinese case, unlike the English case, I produce the answers by manipulating uninterpreted formal symbols. As far as the Chinese is concerned, I simply behave like a computer; I perform computational operations on formally specified elements. For the purposes of the Chinese, I am simply an instantiation of the computer program."

In other words, Searle is proposing that if a computer is just an arrangement of steel balls or electric charges then its content is meaningless without some other phenomenon. Block (1978) used the analogy of a system composed of the population of China communicating with each other to suggest the same idea, that an arrangement of identical things has no meaningful content without a conscious observer who understands its form.

Searle's objection does not convince Direct Realists because they would maintain that 'meaning' is only to be found in objects of perception.

The meaning of meaning and the Symbol Grounding Problem

In his Chinese Room Argument Searle shows that symbols on their own do not have any meaning. In other words, a computer that is a set of electrical charges or flowing steel balls is just a set of steel balls or electrical charges. Leibniz spotted this problem in the seventeeth century.

Searle's argument is also, partly, the **Symbol Grounding Problem**; Harnad (2001) defines this as:

"the symbol grounding problem concerns how the meanings of the symbols in a system can be grounded (in something other than just more ungrounded symbols) so they can have meaning independently of any external interpreter."

Harnad defines a Total Turing Test in which a robot connected to the world by sensors and actions might be judged to be indistinguishable from a human being. He considers that a robot that passed such a test would overcome the symbol grounding problem. Unfortunately Harnad does not tackle Leibniz's misgivings about the internal state of the robot being just a set of symbols (cogs and wheels/charges etc.). The Total Turing Test is also doubtful if analysed in terms of information systems alone, for instance, Powers (2001) argues that an

information system could be grounded in Harnad's sense if it were embedded in a virtual reality rather than the world around it.

So what is "meaning" in an information system? In information systems a **relation** is defined in terms of what thing contains another thing. Having established that one thing contains another this thing is called an **attribute**. A car contains seats so seats are an attribute of cars. Cars are sometimes red so cars sometimes have the attribute "red". This containing of one thing by another leads to classification hierarchies known as a relational database. What Harnad was seeking to achieve was a connection between items in the database and items in the world outside the database. This did not succeed in giving "meaning" to the signals within the machine - they were still a set of separate signals in a materialist model universe.

Aristotle and Plato had a clear idea of meaning when they proposed that ideas depend upon internal images or forms. Plato, in particular conceived that understanding is due to the forms in phenomenal consciousness. Bringing this view up to date, this implies that the way one form contains another gives us understanding. The form of a car contains the form we call seats etc. Even things that we consider to be "content" rather than "form", such as redness, require an extension in space so that there is a red area rather than red by itself (cf: Hume 1739). So if the empiricists are correct our minds contain a geometrical classification system ("what contains what") or geometrical relational database.

A geometrical database has advantages over a sequential database because items within it are highly classified (their relations to other items being implicit in the geometry) and can also be easily related to the physical position of the organism in the world. It would appear that the way forward for artificial consciousness would be to create a virtual reality within the machine. Perhaps the brain works in this fashion and dreams, imagination and hallucinations are evidence for this. But although this would be closer to our experience it still leaves us with the Hard Problem of how the state of a model could become conscious experience.

- Harnad, S. (2001). Grounding Symbols in the Analog World With Neural Nets -- a Hybrid Model, Psycoloquy: 12,#34 http://psycprints.ecs.soton.ac.uk/archive/00000163/#html

- Powers, D.M.W. (2001) A Grounding of Definition, Psycoloquy: 12,#56 http://psycprints.ecs.soton.ac.uk/archive/00000185/#html

Artificial consciousness beyond information processing

The debate about whether a machine could be conscious under any circumstances is usually described as the conflict between physicalism and dualism. Dualists believe that there is something non-physical about consciousness whilst physicalists hold that all things are physical.

Physicalists are not limited to those who hold that consciousness is a property of encoded information on carrier signals. Several indirect realist philosophers and scientists have proposed that, although information processing might deliver the content of consciousness, the state that is consciousness is due to some other physical phenomenon. The eminent neurologist Wilder Penfield was of this opinion and scientists such as Arthur Stanley Eddington, Roger Penrose, Herman Weyl, Karl Pribram and Henry Stapp amongst many

others have also proposed that consciousness involves physical phenomena subtler than information processing. Even some of the most ardent supporters of consciousness in information processors such as Dennett suggest that some new, emergent, scientific theory may be required to account for consciousness.

As was mentioned above, neither the ideas that involve direct perception nor those that involve models of the world in the brain seem to be compatible with current physical theory. It seems that new physical theory may be required and the possibility of dualism is not, as yet, ruled out.

The Computability Problem and Halting of Turing Machines

The Church-Turing thesis

In computability theory the Churchâ€“Turing thesis, Church's thesis, Church's conjecture or Turing's thesis, named after Alonzo Church and Alan Turing, is a hypothesis about the nature of mechanical calculation devices, such as electronic computers. The thesis claims that any calculation that is possible can be performed by an algorithm running on a computer, provided that sufficient time and storage space are available.

This thesis, coupled with the proposition that all computers can be modelled by Turing Machines, means that Functionalist theories of consciousness are equivalent to the hypothesis that the brain operates as a Turing Machine.

Turing machines

A Turing Machine is a pushdown automaton made more powerful by relaxing the last-in-first-out requirement of its stack. (Interestingly, this seemingly minor relaxation enables the Turing machine to perform such a wide variety of computations that it can serve as a model for the computational capabilities of all modern computer software.)

A Turing machine can be constructed using a single tape. There is no requirement for data to be arranged congruently with input or output data so a two dimensional square in the world would be handled as a string or set of strings in the machine yet still calculate a known function. This is problematic in consciousness studies because phenomenal consciousness has many things simultaneously present in several directions at an instant and this form is not congruent with a one dimensional tape.

A Turing machine consists of:

8. A *tape* which is divided into cells, one next to the other. Each cell contains a symbol from some finite alphabet. The alphabet contains a special *blank* symbol (here written as '0') and one or more other symbols. The tape is assumed to be arbitrarily extendible to the left and to the right, i.e., the Turing machine is always supplied with as much tape as it needs for its computation. Cells that have not been written to before are assumed to be filled with the blank symbol.

9. A *head* that can read and write symbols on the tape and move left and right.

10. A *state register* that stores the state of the Turing machine. The number of different states is always finite and there is one special *start state* with which the state register is

initialized.

11. An *action table* (or *transition function*) that tells the machine what symbol to write, how to move the head ('L' for one step left, and 'R' for one step right) and what its new state will be, given the symbol it has just read on the tape and the state it is currently in. If there is no entry in the table for the current combination of symbol and state then the machine will halt.

Note that every part of the machine is finite; it is the potentially unlimited amount of tape that gives it an unbounded amount of storage space.

Another problem arises with Turing Machines is that some algorithms can be shown to be undecidable and so the machine will never halt.

The halting problem

The proof of the halting problem proceeds by reductio ad absurdum. We will assume that there is an algorithm described by the function halt(a, i) that decides if the algorithm encoded by the string *a* will halt when given as input the string *i*, and then show that this leads to a contradiction.

We start with assuming that there is a function halt(a, i) that returns **true** if the algorithm represented by the string *a* halts when given as input the string *i*, and returns **false** otherwise. (The existence of the universal Turing machine proves that every possible algorithm corresponds to at least one such string.) Given this algorithm we can construct another algorithm trouble(s) as follows:

```
function trouble(string s)
   if halt(s, s) = false
      return true
   else
      loop forever
```

This algorithm takes a string *s* as its argument and runs the algorithm halt, giving it *s* both as the description of the algorithm to check and as the initial data to feed to that algorithm. If halt returns **false**, then trouble returns **true**, otherwise trouble goes into an infinite loop. Since all algorithms can be represented by strings, there is a string *t* that represents the algorithm trouble. We can now ask the following question:

Does trouble(t) halt?

Let us consider both possible cases:

1. Assume that trouble(t) halts. The only way this can happen is that halt(t, t) returns **false**, but that in turn indicates that trouble(t) does not halt. Contradiction.

2. Assume that trouble(t) does not halt. Since halt always halts, this can only happen when trouble goes into its infinite loop. This means that halt(t, t) must have returned **true**, since trouble would have returned immediately if it returned **false**. But that in turn would mean that trouble(t) does halt. Contradiction.

Since both cases lead to a contradiction, the initial assumption that the algorithm halt exists must be false.

This classic proof is typically referred to as the **diagonalization proof**, so called because if one imagines a grid containing all the values of halt(a, i), with every possible *a* value given its own row, and every possible *i* value given its own column, then the values of halt(s, s) are arranged along the main diagonal of this grid. The proof can be framed in the form of the question: what row of the grid corresponds to the string *t*? The answer is that the trouble function is devised such that halt(t, i) differs from every row in the grid in at least one position: namely, the main diagonal, where *t=i*. This contradicts the requirement that the grid contains a row for every possible *a* value, and therefore constitutes a proof by contradiction that the halting problem is undecidable.

The simulation argument

According to this argument (Bostrom 2003) the universe could be a giant computer simulation that contains people as well as objects. Bostrom seems to believe that at any instant a collection of bits of information like electrons on silicon or specks of dust on a sheet could be conscious, he states that:

"A common assumption in the philosophy of mind is that of substrate-independence. The idea is that mental states can supervene on any of a broad class of physical substrates. Provided a system implements the right sort of computational structures and processes, it can be associated with conscious experiences."

He then goes on to argue that because of this assumption human beings could be simulations in a computer. Unfortunately, without tackling the problem of how a pattern of dust at an instant could be a person with 'conscious experience' the simulation argument is flawed. In fact even a person made of a moving pattern of dust over several instants is problematical without the assumptions of naive realism or dualism. Bostrom, in evoking supervenience is probably a dualist; he puts 'mental' states' beyond physical explanation (ie: simply assumes that conscious mental states could exist in a pattern of electrons, dust or steel balls etc.). In view of this dualism, Bostrom's argument reduces to the proposal that the world is a digital simulation apart from something else required for endowing the simulations of people in the world with consciousness.

Notes and References

Note 1: Strictly this is the quantum 'amplitude' for the electron to go in a particular direction rather than the probability.

The philosophical problem

- Chalmers, D. (1996). The Conscious Mind. New York: Oxford University Press.

Epiphenomenalism and the problem of change

- Huxley, T. H. (1874.) On the Hypothesis that Animals are Automata, and its History. The Fortnightly Review: 16:555-580.

The Problem of Time

- Atmanspacher, H. (1989). The aspect of information production in the process of observation, in: Foundations of Physics, vol. 19, 1989, pp. 553-77

- Atmanspacher, H. (2000).Ontic and epistemic descriptions of chaotic systems. In Proceedings of CASYS 99, ed. by D. Dubois, Springer, Berlin 2000, pp. 465-478. http://www.igpp.de/english/tda/pdf/liege.pdf

- de Broglie, L. (1925) On the theory of quanta. A translation of : RECHERCHES SUR LA THEORIE DES QUANTA (Ann. de Phys., 10e sÂ´erie, t. III (Janvier-F Â´evrier 1925).by: A. F. Kracklauer. http://www.nonloco-physics.000freehosting.com/ldb_the.pdf

- Brown, K. (????) Mathpages 3.7 Zeno and the Paradox of Motion. http://www.mathpages.com/rr/s3-07/3-07.htm

- Brown, K. (????) Mathpages Zeno and Uncertainty. http://www.mathpages.com/home/kmath158.htm

- Franck, G. (1994). Published in: Harald Atmanspacher and Gerhard J. Dalenoort (eds.), Inside Versus Outside. Endo- and Exo-Concepts of Observation and Knowledge in Physics, Philosophy, and Cognitive Science, Berlin: Springer,1994, pp. 63-83 http://www.iemar.tuwien.ac.at/publications/GF_1994a.pdf

- Lynds, P. (2003). Time and Classical and Quantum Mechanics: Indeterminacy vs. Discontinuity. Foundations of Physics Letters, 16(4), 2003. http://doc.cern.ch//archive/electronic/other/ext/ext-2003-042.pdf

- McCall, S. 2000. QM and STR. The combining of quantum mechanics and relativity theory. Philosophy of Science 67 (Proceedings), pp. S535-S548. http://www.mcgill.ca/philosophy/faculty/mccall/

- McTaggart, J.M.E. (1908) The Unreality of Time. Published in Mind: A Quarterly Review of Psychology and Philosophy 17 (1908): 456-473. http://www.ditext.com/mctaggart/time.html

- Petkov, V. (2002). Montreal Inter-University Seminar on the History and Philosophy of Science. http://alcor.concordia.ca/~vpetkov/absolute.html

- Pollock, S. (2004) Physics 2170 - Notes for section 4. University of Colorado. http://www.colorado.edu/physics/phys2170/phys2170_spring96/notes/2170_notes4_18.html

- Weyl, H. (1920). Space, Time, Matter.(Dover Edition).

Further reading:

- James, W. (1890). The Principles of Psychology. CHAPTER XV. THE PERCEPTION OF TIME.

- Ellis McTaggart, J.M. (1908) The Unreality of Time. Mind: A Quarterly Review of Psychology and Philosophy 17 (1908): 456-473.

- McKinnon, N.(2003)Presentism and Consciousness. Australasian Journal of Philosophy 81:3 (2003), 305-323.

- Lamb, A.W. (1998) Granting Time Its Passage. Twentieth World Congress of Philosophy Boston, Massachusetts U.S.A. 10-15 August 1998

- Franck, G. (1994). Physical Time and Intrinsic Temporality. Published in: Harald Atmanspacher and Gerhard J. Dalenoort (eds.), Inside Versus Outside. Endo- and Exo-Concepts of Observation and Knowledge in Physics, Philosophy, and Cognitive Science, Berlin: Springer,1994, pp. 63-83

- Lynds, P. (2003). Subjective Perception of Time and a Progressive Present Moment: The Neurobiological Key to Unlocking Consciousness.

- Alfred North Whitehead. (1920) "Time". Chapter 3 in The Concept of Nature. Cambridge: Cambridge University Press (1920): 49-73.

- Franck, G. HOW TIME PASSES. On Conceiving Time as a Process. Published in : R. Buccheri/ M. Saniga/ W.M. Stuckey (eds.), The Nature of Time: Geometry, Physics and Perception (NATO Science Series), Dodrecht: Kluwer Academic, 2003, pp. 91-103

- Savitt, S.F. (1998). There's no time like the present (in Minkowski space-time).

- Le Poidevin, R. (2004) The Experience and Perception of Time. Stanford Encyclopedia of Philosophy.

- Norton, J. (2004) The Hole Argument. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/spacetime-holearg/index.html

- Rovelli, C. (2003) Quantum Gravity. Book. http://www.cpt.univ-mrs.fr/~rovelli/book.pdf

- Penrose, R. 1989. The Emperor's New Mind: Concerning Computers, Minds, and Laws of Physics. New York and Oxford: Oxford University Press

- Stein, H. 1968. On Einstein-Minkowski Space-Time, The Journal of Philosophy 65: 5-23.

- Torretti, R. 1983. Relativity and Geometry. Oxford, New York, Toronto, Sydney, Paris, Frankfurt: Pergamon Press.

The existence of time

- Clay, ER (1882). The Alternative: A Study in Psychology, p. 167. (Quoted in James 1890).

- Gombrich, Ernst (1964) 'Moment and Movement in Art', Journal of the Warburg and Courtauld Institutes XXVII, 293-306. Quoted in Le Poidevin, R. (2000). The Experience and Perception of Time. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/time-experience

- James, W. (1890) .The Principles of Psychology http://psychclassics.yorku.ca/James/Principles/prin15.htm

- Lindner et al (2005) Attosecond double-slit experiment. Accepted for Physical Review Letters. http://arxiv.org/abs/quant-ph/0503165

- Paulus, GG et al. (2003) PRL 91, 253004 (2003), http://mste.laser.physik.uni-muenchen.de/paulus.pdf

- Physics Web. New look for classic experiment. http://physicsweb.org/articles/news/9/3/1/1?rss=2.0

- Rea, MC. (2004). Four Dimensionalism. The Oxford Handbook for Metaphysics http://www.nd.edu/~mrea/Online%20Papers/Four%20Dimensionalism.pdf

- Romer, H. (2004) Weak Quantum Theory and the Emergence of Time http://arxiv.org/PS_cache/quant-ph/pdf/0402/0402011.pdf

- Amjorn, J et al (2004). Emergence of a 4D world from causal quantum gravity. Phys.Rev.Lett. 93 (2004) 131301 http://www.arxiv.org/PS_cache/hep-th/pdf/0404/0404156.pdf

Useful Links

- The web site of Dr Paulus, one of the principle physicists working on these femtosecond laser projects. http://faculty.physics.tamu.edu/ggp/

Relationalism, Substantivalism etc..

- Earman, J. (2002). Thoroughly Modern McTaggart. Philosophers' Imprint. Vol. 2 No. 3. August 2002. http://www.umich.edu/~philos/Imprint/frameset.html?002003+28+pdf

- Einstein, A. (1916b). Die Grundlage der allgemeinin Relativitatstheorie, Annalen der Physik. 49, 769 (1916); translated by W.Perrett and G.B.Je The Foundations of of the tivity General Theory of Relativity, in The Principle of Relativity (Dover, New York, 1952), pp.117-118. Pointed out by Lusanna and Pauri in their draft of "General Covariance and the Objectivity of Space-Time Point Events".

- Gardner, M. (1990). The New Ambidextrous Universe: Symmetry and Asymmetry, from Mirror Reflections to Superstrings.WH Freeman & Co. New York.

- Gaul, M. & Rovelli, C. (1999). Loop Quantum Gravity and the Meaning of Diffeomorphism Invariance. http://arxiv.org/PS_cache/gr-qc/pdf/9910/9910079.pdf

- MacDonald, A. (2001). Einstein's Hole Argument. Am. J. Phys. 69, 223-225 (2001). http://faculty.luther.edu/~macdonal/HoleArgument.pdf

- Norton, J.D. (1993). General covariance and the foundations of general relativity: eight decades of dispute. Rep. Prog. Phys. 56 (1993) 791-858. http://www.pitt.edu/~jdnorton/papers/decades.pdf

- Norton, J.D. (1999) A Conjecture on Einstein, the Independent Reality of Spacetime Coordinate Systems and the Disaster of 1913. http://philoscience.unibe.ch/lehre/sommer05/Einstein%201905/Texte/113

- Pooley, O. (2002). Handedness, parity violation,To appear in Katherine Brading and Elena Castellani (eds), in preparation, Symmetries in Physics: Philosophical Reflections (Cambridge: Cambridge University Press). http://users.ox.ac.uk/~ball0402/papers/parity.pdf

Quantum theory and time

- Hagan, S., Hammeroff, S.R. and Tuszynski, J.A.(2002). Quantum computation in brain microtubules: Decoherence and biological feasibility. PHYSICAL REVIEW E, VOLUME 65, 061901. http://arxiv.org/abs/quant-ph/0005025

- Hawking, S. (1999) The future of quantum cosmology. http://www.hawking.org.uk/ps/futquan.ps

- Isham, C.J. (1993). Canonical quantum gravity and the problem of time.In Integrable Systems, Quantum Groups, and Quantum Field heories, pages 157-288. Kluwer Academic Publishers, London, 1993. http://arxiv.org/PS_cache/gr-qc/pdf/9210/9210011.pdf

- Isham, C.J. Structural Issues in Quantum Gravity. http://lanl.arxiv.org/PS_cache/gr-qc/pdf/9510/9510063.pdf

- Jacobson, T. (1995). Thermodynamics of Spacetime: The Einstein Equation of State. Phys.Rev.Lett. 75 (1995) 1260-1263 http://lanl.arxiv.org/PS_cache/gr-qc/pdf/9504/9504004.pdf

- Tegmark, M. (2000). The Importance of Quantum Decoherence in Brain Processes. Phys.Rev. E61 (2000) 4194-4206 http://arxiv.org/PS_cache/quant-ph/pdf/9907/9907009.pdf

- Zeh, D. (2001) The Physical basis of the direction of time. Fourth edition (ISBN 3-540-42081-9)- Springer-Verlag http://www.rzuser.uni-heidelberg.de/~as3/time-direction/

The problem of qualia

- Alter, T. (1998). "A Limited Defense of the Knowledge Argument." Philosophical Studies 90: 35-56. But especially the discussion at the following web site: http://host.uniroma3.it/progetti/kant/field/ka.html

- Anglin, J.R. & Zurek, J.H. (1996). Decoherence of quantum fields: decoherence and predictability. Phys.Rev. D53 (1996) 7327-7335 http://arxiv.org/PS_cache/quant-ph/pdf/9510/9510021.pdf

- Bacciagaluppi, G. (2004). The role of decoherence in quantum theory. http://plato.stanford.edu/entries/qm-decoherence/

- Dennett, D. (1991), Consciousness Explained, Boston: Little Brown and Company

- Dretske, F. (2003). Experience as Representation. Philosophical Issues 13, 67-82. http://humanities.ucsc.edu/NEH/dretske1.htm

- Jackson, F. (1982) Epiphenomenal Qualia. Philosophical Quarterly, 32 (1982), pp. 127-36. http://instruct.westvalley.edu/lafave/epiphenomenal_qualia.html

- Lehar, S. (2003) Gestalt Isomorphism and the Primacy of the Subjective Conscious Experience: A Gestalt Bubble Model. (2003) Behavioral & Brain Sciences 26(4), 375-444. http://cns-alumni.bu.edu/~slehar/webstuff/bubw3/bubw3.html

- Levine, J. (1983) 'Materialism and Qualia: The Explanatory Gap', Pacific. Philosophical Quarterly, 64: 354-61.

- Lycan, W. (1987). Consciousness, Cambridge, Mass : The MIT Press.
- Ogborn, J. & Taylor, E.F. (2005) Quantum physics explains Newton's laws of motion. Physics Education 40(1). 26-34. http://www.eftaylor.com/pub/OgbornTaylor.pdf
- Strawson, G. (1994). Mental Reality, Cambridge USA: the MIT Press, Bradford Books.
- Tye, M. (1995). Ten Problems of Consciousness (Bradley Books, MIT Press),
- Tye, M. (2003). Visual qualia and visual content revisited. Ed. David Chalmers. OUP. http://sun.soci.niu.edu/~phildept/MT/Visual.pdf
- Tye, M. (2003). Qualia. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/qualia/
- Zurek, W.H. (2003). Decoherence, einselection and the quantum origins of the classical. Rev. Mod. Phys. 75, 715 (2003) http://arxiv.org/PS_cache/quant-ph/pdf/0105/0105127.pdf

Machine and digital consciousness

- Block, N. (1978). "Trouble with functionalism", In W. Savage (ed.),Perception and Cognition: Minnesota Studies in Philosophy of Science, Vol IX, Minnesota University Press, 1978, pp. 261-362; reprinted in Block (ed.) (1980), vol. I, pp. 268-305; reprinted (excerpt) in Lycan (ed.)(1990), pp. 444-468.
- Searle, J.R. 1980. Minds Brains and Programs. The Behavioral and Brain Sciences, vol. 3. Copyright 1980 Cambridge University Press.
- Bostrom, N. 2003. Are you living in a Computer Simulation? Philosophical Quarterly, 2003, Vol. 53, No. 211, pp. 243-255.

The Measurement Problem

In quantum physics the probability of an event is deduced by taking the square of the **amplitude** for an event to happen. The term "amplitude for an event" arises because of the way that the Schrodinger equation is derived using the mathematics of ordinary, classical waves where the amplitude over a small area is related to the number of photons hitting the area. In the case of light, the probability of a photon hitting that area will be related to the ratio of the number of photons hitting the area divided by the total number of photons released. The number of photons hitting an area per second is the intensity or amplitude of the light on the area, hence the probability of finding a photon is related to "amplitude".

However, the Schrodinger equation is not a classical wave equation. It does not determine events, it simply tells us the probability of an event. In fact the Schrodinger equation in itself does not tell us that an event occurs at all, it is only when a measurement is made that an event occurs. The measurement is said to cause *state vector reduction*. This role of measurement in quantum theory is known as the **measurement problem**. The measurement problem asks how a definite event can arise out of a theory that only predicts a continuous probability for events.

Two broad classes of theory have been advanced to explain the measurement problem. In the first it is proposed that observation produces a sudden change in the quantum system so that a particle becomes localised or has a definite momentum. This type of explanation is known as *collapse of the wavefunction*. In the second it is proposed that the probabilistic Schrodinger equation is always correct and that, for some reason, the observer only observes one particular outcome for an event. This type of explanation is known as the *relative state interpretation*. In the past thirty years relative state interpretations, especially Everett's relative state interpretation have become favoured amongst quantum physicists.

The quantum probability problem

The measurement problem is particularly problematical when a single particle is considered. Quantum theory differs from classical theory because it is found that a single photon seems to be able to interfere with itself. If there are many photons then probabilities can be expressed in terms of the ratio of the number hitting a particular place to the total number released but if there is only one photon then this does not make sense. When only one photon is released from a light source quantum theory still gives us a probability for a photon to hit a particular area but what does this mean at any instant if there is indeed only one photon?

If the Everettian interpretation of quantum mechanics is invoked then it might seem that the probability of the photon hitting an area in your particular universe is related to the occurrences of the photon in all the other universes. But in the Everrettian interpretation even the improbable universes occur. This leads to a problem known as the **quantum probability problem**:

> If the universe splits after a measurement, with every possible measurement outcome realised in some branch, then how can it make sense to talk about the probabilities of each outcome? Each outcome occurs.

This means that if our phenomenal consciousness is a set of events then there would be

endless copies of these sets of events, almost all of which are almost entirely improbable to an observer outside the brain but all of which exist according to an Everrettian Interpretation. Which set is you? Why should 'you' conform to what happens in the environment around you?

The preferred basis problem

It could be held that you assess probabilities in terms of the branch of the universe in which you find yourself but then why do you find yourself in a particular branch? Decoherence Theory is one approach to these questions. In decoherence theory the environment is a complex form that can only interact with particles in particular ways. As a result quantum phenomena are rapidly smoothed out in a series of micro-measurements so that the macro-scale universe appears quasi-classical. The form of the environment is known as the preferred basis for quantum decoherence. This then leads to the **preferred basis problem** in which it is asked how the environment occurs or whether the state of the environment depends on any other system.

According to most forms of decoherence theory 'you' are a part of the environment and hence determined by the preferred basis. From the viewpoint of phenomenal consciousness this does not seem unreasonable because it has always been understood that the conscious observer does not observe things as quantum superpositions. The conscious observation is a classical observation.

However, the arguments that are used to derive this idea of the classical, conscious observer contain dubious assumptions that may be hindering the progress of quantum physics. The assumption that the conscious observer is simply an information system is particularly dubious:

"Here we are using aware in a down - to - earth sense: Quite simply, observers know what they know. Their information processing machinery (that must underlie higher functions of the mind such as "consciousness") can readily consult the content of their memory. (Zurek 2003).

This assumption is the same as assuming that the conscious observer is a set of measurements rather than an observation. It makes the rest of Zurek's argument about decoherence and the observer into a tautology - given that observations are measurements then observations will be like measurements. However, conscious observation is not simply a change of state in a neuron, a "measurement", it is the entire manifold of conscious experience.

In his 2003 review of this topic Zurek makes clear an important feature of information theory when he states that:

"There is no information without representation."

So the contents of conscious observation are states that correspond to states of the environment in the brain (ie: measurements). But how do these states in the brain arise? The issue that arises here is whether the representation, the contents of consciousness, is entirely due to the environment or due to some degree to the form of conscious observation. Suppose we make the reasonable assumption that conscious observation is due to some physical field in the dendrites of neurons rather than in the action potentials that transmit the

state of the neurons from place to place. This field would not necessarily be constrained by decoherence; there are many possibilities for the field, for instance, it could be a radio frequency field due to impulses or an electromagnetic field (cf: Anglin & Zurek (1996)) or some quantum state of macromolecules etc.. Such a field might contain many superposed possibilities for the state of the underlying neurons and although these would not affect sensations, they could affect the firing patterns of neurons and create actions in the world that are not determined by the environmental "preferred basis".

Zeh (2000) provides a mature review of the problem of conscious observation. For example he realises that memory is not the same as consciousness:

"The genuine carriers of consciousness ... must not in general be expected to represent memory states, as there do not seem to be permanent contents of consciousness."

and notes of memory states that they must enter some other system to become part of observation:

"To most of these states, however, the true physical carrier of consciousness somewhere in the brain may still represent an external observer system, with whom they have to interact in order to be perceived. Regardless of whether the ultimate observer systems are quasi-classical or possess essential quantum aspects, consciousness can only be related to factor states (of systems assumed to be localized in the brain) that appear in branches (robust components) of the global wave function - provided the Schrodinger equation is exact. Environmental decoherence represents entanglement (but not any "distortion" of the brain, in this case), while ensembles of wave functions, representing various potential (unpredictable) outcomes, would require a dynamical collapse (that has never been observed)."

However, Zeh (2003) points out that events may be irreversibly determined by decoherence before information from them reaches the observer. This might give rise to a multiple worlds and multiple minds mixture for the universe, the multiple minds being superposed states of the part of the world that is the mind. Such an interpretation would be consistent with the *apparently* epiphenomenal nature of mind. A mind that interacts only weakly with the consensus physical world, perhaps only approving or rejecting passing actions would be an ideal candidate for a QM multiple minds hypothesis.

Further reading and references

- Anglin, J.R. & Zurek, J.H. (1996). Decoherence of quantum fields: decoherence and predictability. Phys.Rev. D53 (1996) 7327-7335 http://arxiv.org/PS_cache/quant-ph/pdf/9510/9510021.pdf

- Baker, D.J. (2004). Lingering Problems with Probability in Everettian Quantum Mechanics http://www.princeton.edu/~hhalvors/teaching/phi538_f2004/LingeringProbsEverett.pdf

- Lockwood, M. (1996) Many Minds Interpretations of quantum mechanics. The British Journal of the Philosophy of Science. 47: 2 (159-188). http://www.ibiblio.org/weidai/Many_Minds.pdf

- Pearl, P. (1997). True collapse and false collapse. Published in Quantum Classical Correspondence: Proceedings of the 4th Drexel Symposium on Quantum Nonintegrability,

Philadelphia, PA, USA, September 8-11, 1994, pp. 51-68. Edited by Da Hsuan Feng and Bei Lok Hu. Cambridge, MA: International Press, 1997. http://arxiv.org/PS_cache/quant-ph/pdf/9805/9805049.pdf

- Squires, E.J. (1996). What are quantum theorists doing at a conference on consciousness? http://arxiv.org/PS_cache/quant-ph/pdf/9602/9602006.pdf

- Zeh, H. D. (1979). Quantum Theory and Time Assymetry. Foundations of Physics, Vol 9, pp 803-818 (1979). http://xxx.lanl.gov/PS_cache/quant-ph/pdf/0307/0307013.pdf

- Zeh, H.D. (2000) THE PROBLEM OF CONSCIOUS OBSERVATION IN QUANTUM MECHANICAL DESCRIPTION. Epistemological Letters of the Ferdinand-Gonseth Association in Biel (Switzerland) Letter No 63.0.1981, updated 2000. http://arxiv.org/PS_cache/quant-ph/pdf/9908/9908084.pdf

- Zeh, H.D. (2003). Decoherence and the Appearance of a Classical World in Quantum Theory, second edition, Authors:. E. Joos, H.D. Zeh, C. Kiefer D. Giulini, J. Kupsch, and I.-O. Stamatescu. Chapter 2: Basic Concepts and their Interpretation. http://www.rzuser.uni-heidelberg.de/~as3/index.html

- Zurek, W.H. (2003). Decoherence, einselection and the quantum origins of the classical. Rev. Mod. Phys. 75, 715 (2003) http://arxiv.org/PS_cache/quant-ph/pdf/0105/0105127.pdf

Part III: The Neuroscience of Consciousness

"All parts of the brain may well be involved in normal conscious processes but the indispensable substratum of consciousness lies outside the cerebral cortex, probably in the diencephalon" Penfield 1937.

"The brain stem-thalamocortical axis supports the state, but not the detailed contents of consciousness, which are produced by cortex" Baars et al 1998.

Introduction

It is recommended that readers review "The philosophical problem" before reading the sections on the neuroscience of consciousness.

One of the most exciting discoveries of neuroscience is that nearly all of the brain performs functions that are not part of conscious experience. In everyday life we are usually unaware of breathing or heartbeats yet there are parts of the brain dedicated to these functions. When we pick up a pencil we have no experience of the fine control of individual muscles yet large areas of cortex and cerebellum implement this. Things do not appear as greyscale and then have the colour poured into them although this strange colour addition is done in the visual cortex. Most of the brain is non-conscious but how is the "ghost in the machine", the mind, created by and linked into the non-conscious brain?

Although most of the processes in the brain are non-conscious there can be little doubt that the output of sensory processes contribute to experience. For example, although we do not experience the process of adding colour to visual data in cortical area V4 we do experience coloured forms and although we have little inkling of the hugely complex creation of words in the temporal/frontal lobes we do experience verbal thoughts. Our experience is an integrated output of most of the brain processes that deal with sensation as well as dreams, thoughts and emotions. But how and where does this experience occur?

The substrate of experience

Quantum mechanical events in the world give rise to signals that travel from these events to the sense organs and the brain. The signals have a state that is related to the properties of the original QM event. Some of these signals form phenomenal consciousness and some are used in the processes of access consciousness. According to neuroscience the signals are physical things such as electromagnetic fields, distributions of chemicals, electrical impulses etc.

The signals used in access consciousness are used in the processes that mediate between stimulus and response. These processes have been investigated in depth. The signals within the processes consist of physical stimuli, the electrical impulses in the cell bodies and axons of nerve cells, the electrical fields in the dendrites of nerve cells, and various chemical signals that connect nerve cells.

The signals that compose phenomenal consciousness have not been elucidated. Perhaps the least likely signals for this role are electrical impulses in nerve fibres because they are distributed unevenly in time and space and can even be absent for relatively long periods.

Furthermore, electrical impulses cannot be easily superimposed on one another. There are many other possibilities however, such as: the electrical fields on the dendrites of neurons, the fields of chemicals spreading out from synapses, the radio-frequency emissions of action potentials, events in the microtubules in cells, the depolarisations of glia, the varying fields measured by EEG devices, the quantum superposition of brain states etc...

Phenomenal consciousness could exist in the dendritic field of ten neurons receiving 100,000 synapses or as an oscillation of fields over the whole brain. The substrate of phenomenal consciousness could be staring us in the face as a state of the whole brain or be like a needle in a haystack, lurking in a tiny region of brain, unsuspected and undiscovered.

Given that there is no widely accepted theory of phenomenal consciousness Crick (1994) and Crick and Koch (1998)approached the problem of the location of the substrate of consciousness by proposing that scientists search for the **Neural Correlates of Consciousness**. These neural correlates consist of events in the brain that accompany events in conscious experience.

References:

Crick, F. (1994). The Astonishing Hypothesis. New York: Scribners.

Crick, F. & Koch, C. (1998).Consciousness and Neuroscience. Cerebral Cortex, 8:97-107, 1998 http://www.klab.caltech.edu/~koch/crick-koch-cc-97.html

Neuroanatomy

General layout of the CNS

The Central Nervous System (CNS) consists of the spinal cord, the brain and the retina.

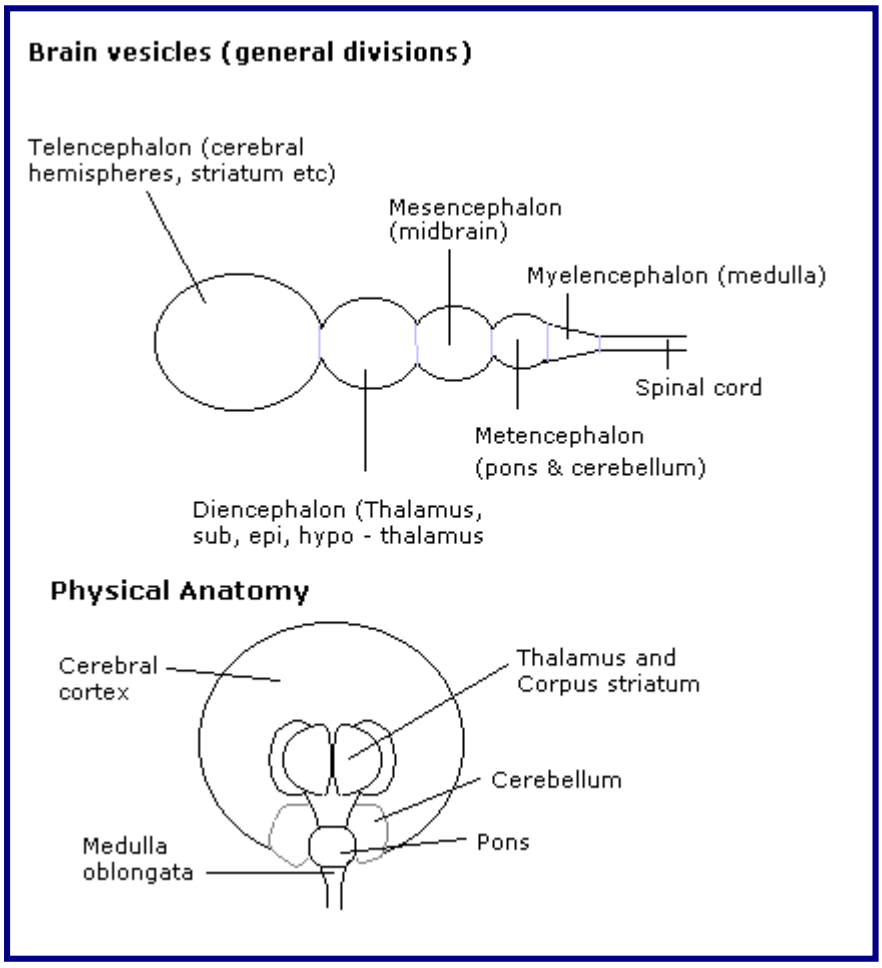
The CNS consists of two major groups of active cells, the **neurons** and the **glia**. The neurons conduct short impulses of electricity along their membranes called '**action potentials** and encode data as frequency modulated signals (ie: different intensities of stimulation are converted into different rates of firing). The glia modify the connections between neurons and can respond to neuron activity by a change of voltage across their membranes. Glia also have many other roles such as sustaining neurons and providing electrical insulation.

Neurons have three principal parts: the **cell body**, the **dendrites** and the **axon**. Impulses flow from the cell body to the axon. The axon can be over a metre long and bundles of axons form **nerve fibres**. Where an axon makes contact with the dendrites or cell body of another neuron there is a special sort of junction called a **synapse**. Transmission of data across synapses is usually mediated by chemical signals.

Areas of the brain where there are many cell bodies have a beige/grey tinge and are called **grey matter**. Areas that contain mainly nerve fibres are called **white matter**. Masses of grey matter outside of the surface of the cerebral cortex or the cerebellum are called **nuclei**.

The brain is of central interest in consciousness studies because consciousness persists even when the spinal cord is sectioned at the neck.

The brain can be divided into five distinct divisions or '*vesicles* on the basis of embryological development. These are the myelencephalon, metencephalon, mesencephalon, diencephalon and telencephalon (See the illustration below).



Myelencephalon: Medulla oblongata.

Metencephalon: pons and cerebellum.

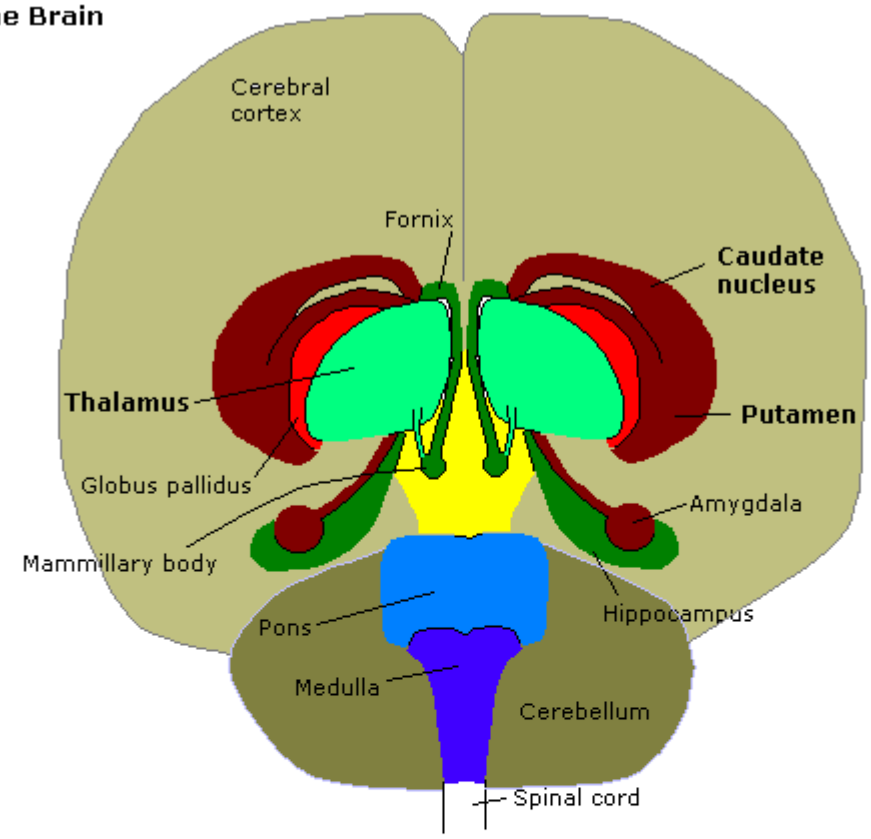
Mesencephalon: midbrain (tectum containing the superior colliculus and inferior colliculus, red nucleus, substantia nigra, cerebellar peduncles.

Diencephalon: thalamus, epithalamus, hypothalamus, subthalamus.

Telencephalon: corpus striatum, cerebral hemispheres.

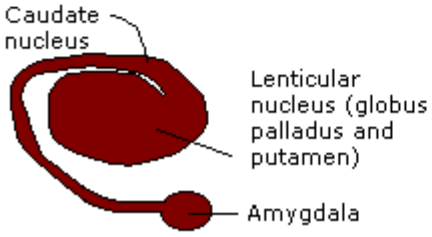
These divisions tend to obscure the physical anatomy of the brain which looks like a rod of spinal cord with a swelling at the top due to the thalamus and corpus striatum. Around the top of the rod is a globe of deeply indented cerebral cortex and at the back there is the puckered mass of cerebellum. The physical anatomy is shown in greater detail in the illustration below where the thalamus and corpus striatum have been splayed out to show more detail.

The Brain



The brain as viewed from the underside and front. The thalamus and Corpus Striatum (Putamen, caudate and amygdala) have been splayed out to show detail.

Corpus Striatum

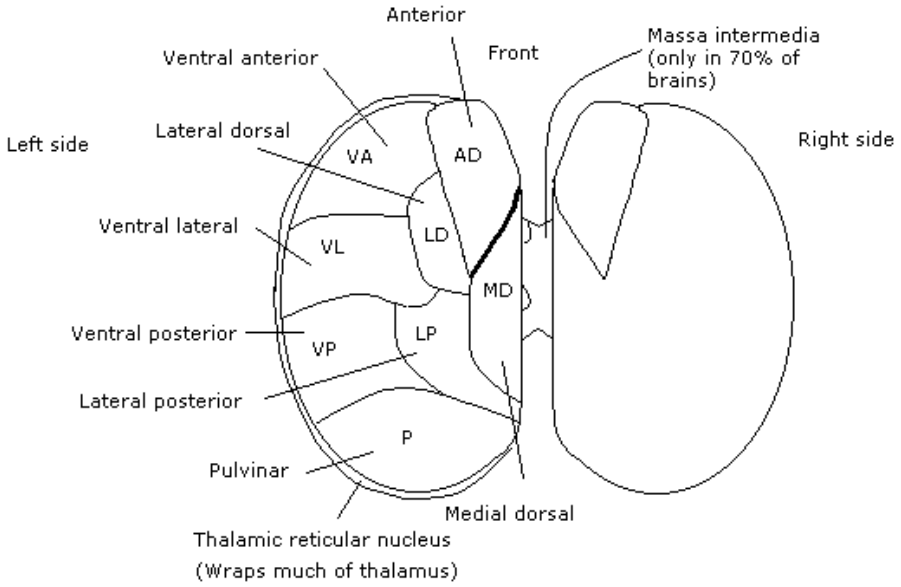


The thalamus is a complex organ with numerous nuclei. These are listed below:

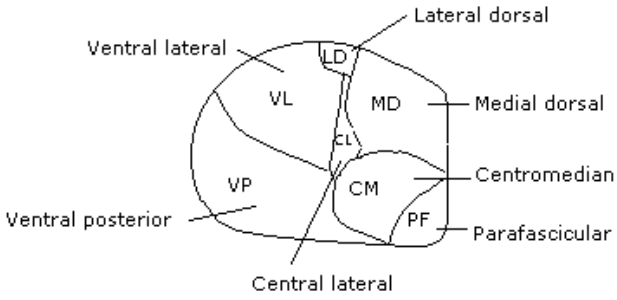
Type of Nucleus	Name	Abbrev	Function
Reticular	Reticular	R	Arousal
Intralaminar	Centromedian	CM	Arousal, attention, motivation, pain
	Parafascicular	Pf	
	Central lateral	CL	
	Paracentral	Pcn	
Intralaminar Midline	Reunions	Re	
	Paraventricular	Pv	
	Rhomboid		
Nonspecific	Pulvinar	P	Association
	Lateral dorsal	LD	
	Anterior	AD	
	Anteromedial	AM	
	Anteroventral	AV	
	Lateral posterior	LP	
	Medial Dorsal	MD	
Specific Thalamic Nuclei	Lateral geniculate	LGN	Vision
(Sensory Relays)	Medial geniculate	MGN	Auditory
	Ventral posterior	VP	General sensation
Specific Thalamic Nuclei	Ventral anterior	VA	Motor
(motor)	Ventral lateral	VL	Motor

The location of these nuclei is shown in the illustration below:

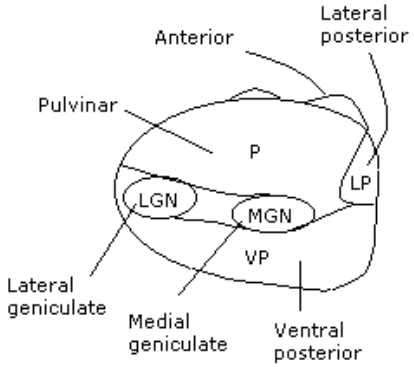
Thalamus Dorsal View (top view).



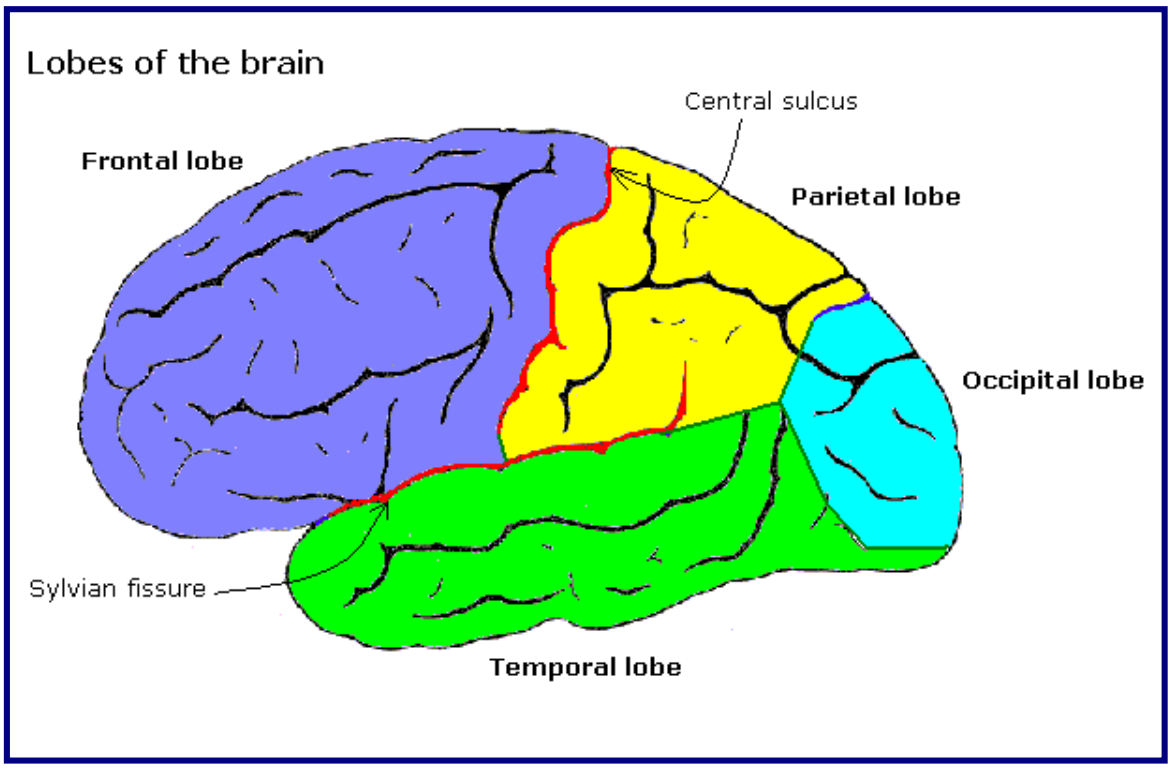
Mid-section



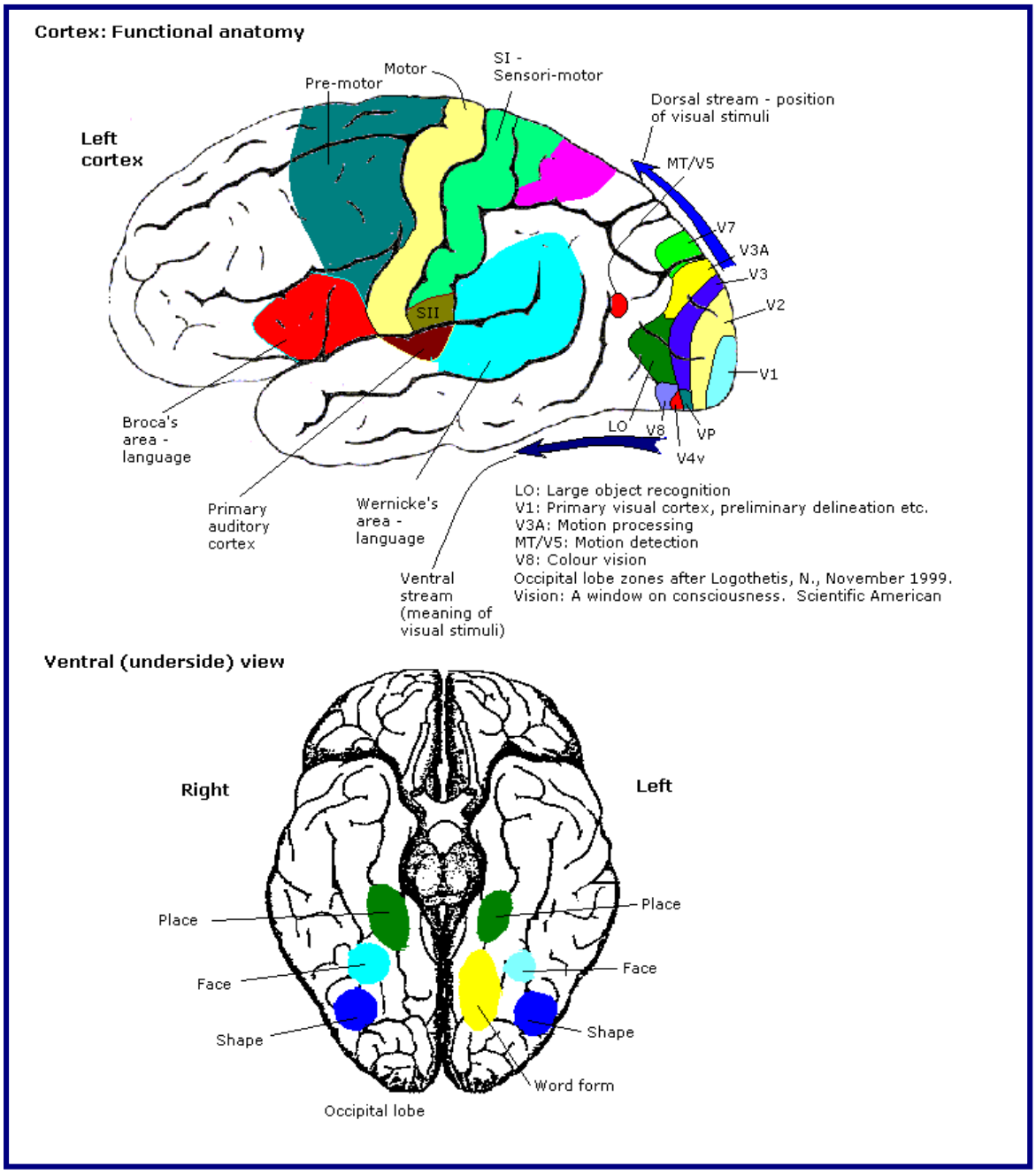
Rear view



The cerebral hemispheres consist of a thin layer of nerve cell bodies on the surface (the cerebral cortex) with a mass of white, interconnecting fibres below (the cerebral medulla). Each hemisphere is divided into four principle **lobes** as shown in the illustration below:



The cortex is a set of interconnected processors. The general layout of the cortex with the location of the processors is shown in the illustration below:



The pathways in the brain tend to preserve the **topography** of the sense organs so that particular groups of cells on the retina, cochlear or body have corresponding groups of cells in the thalamus or cortex. The retina is said to have a topological mapping onto the thalamus so that the projection of the optic nerve is said to be **retinotopic**.

Nerve fibres that go to a part of the brain are called **afferents** and fibres that come from a part of the brain are called **efferents**.

The cortex and thalamus/striatum are intimately linked by millions of connecting fibres and there is also a direct connection from the motor cortex to the spinal cord.

Sensory pathways

Information from the sense organs travels along the appropriate sensory nerve (optic, auditory, spinal etc.) and once in the brain is divided into three principal paths that connect either with the thalamus, the cerebellum or the reticular formation.

There are thalamic nuclei for each broad type of sensation and these have reciprocal connections with specific areas of cortex that deal with the appropriate mode of sensation. The large mass of nerve fibres that mediate the connection between the thalamus and cortex are known as the thalamo-cortical and cortico-thalamic tracts. There tend to be more sensory nerve fibres returning from the cortex to the thalamus than connect from the thalamus to the cortex so it is difficult to determine whether the cortex is the destination of sensory data or a region that supplies extra processing power to thalamic nuclei.

The cerebellum mediates reflex control of complex movements and receives input from most of the sense organs.

The reticular formation is a group of loosely distributed neurons in the medulla, pons and mesencephalon. It receives a large amount of autonomic input and also input from all the sense organs. The intralaminar nuclei of the thalamus are the principal destination of reticular output to higher centres. In the most primitive vertebrates the reticular formation performs most of the higher control functions of the animal. The reticular formation is implicated in the maintenance of sleep-wake cycles and activates the higher centres. This activity has attracted the label **ascending reticular activating system** (ARAS) to describe how the activity of higher centres is controlled by reticular input. This title is unfortunate from the point of view of consciousness studies because it implies that conscious experience is a result of activating the cortex when it could be due to turning on or off particular systems all the way from the reticular formation to the cortex. Destruction of the reticular formation leads to coma.

Motor and output pathways

Motor control of the body below the skull is accomplished by three principle routes.

The motor cortex of the frontal lobes and related cortex in the parietal lobes can control movement directly via nerves known as the cortico-spinal tract (also called the pyramidal tract). The activity of the motor cortex is modified and controlled by a loop that passes through the corpus striatum, the substantia nigra and the subthalamic nucleus and returns to the cortex. These controlling nuclei are, along with the amygdala, known as the **basal**

ganglia.

The cerebellum and the corpus striatum provide complex reflex control of the body through nerves that travel through the red nucleus and form the rubro-spinal tract.

The vestibular nucleus, which processes signals related to balance and posture, has direct connections with the periphery via the vestibulo-spinal tract.

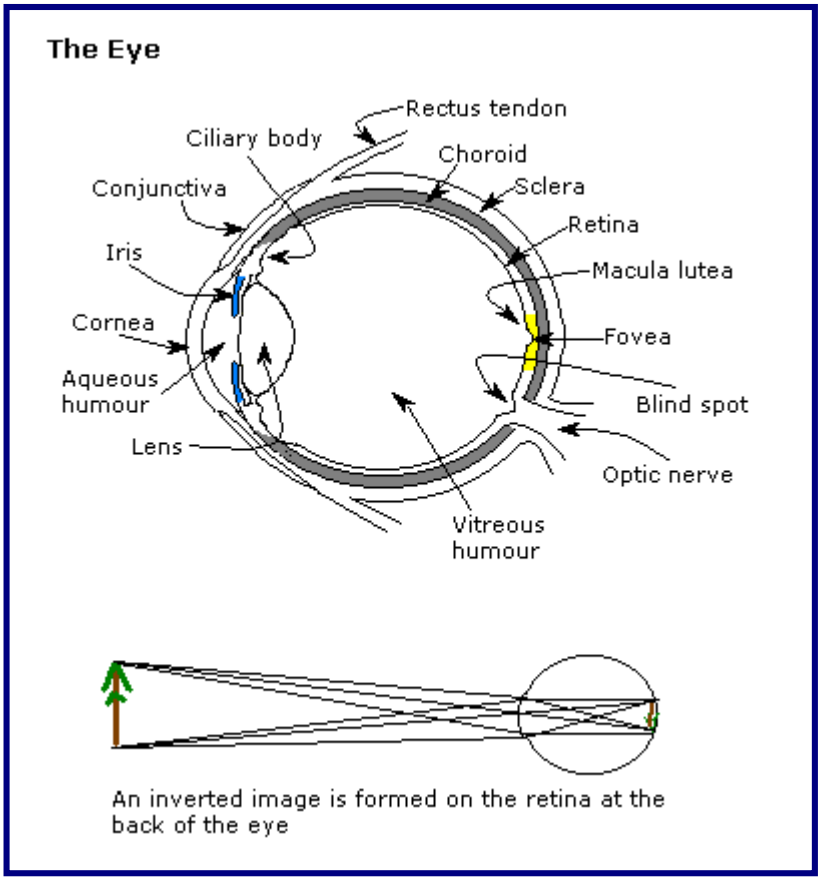
Apart from the routes for controlling motor activity there are also other outputs from the brain, for instance the autonomic nervous system is intimately linked with the reticular formation which has areas that control blood pressure, respiratory rhythm etc.

The neurophysiology of sensation and perception

Vision

The human eye

The eye is a remarkable optical instrument that is often poorly understood by students of consciousness. The most popular misconception is that there is a 'focus' within the eye through which all the light rays pass! The purpose of this article is to describe our knowledge of the optics of the eye so that such misconceptions can be avoided.



The eye consists of several surfaces at which refraction occurs: air-cornea, cornea-aqueous humour, aqueous humour-lens, lens-vitreous humour. The crude image forming capability of the eye can be represented quite accurately by the *reduced eye* model which involves a single optical surface (air-cornea). Optometrists use more accurate models such as the Gullstrand Schematic Eye, the Le Grand Theoretical and the LeGrand Simplified Eye.

The lens system at the front of the eye forms an inverted image on the retina.

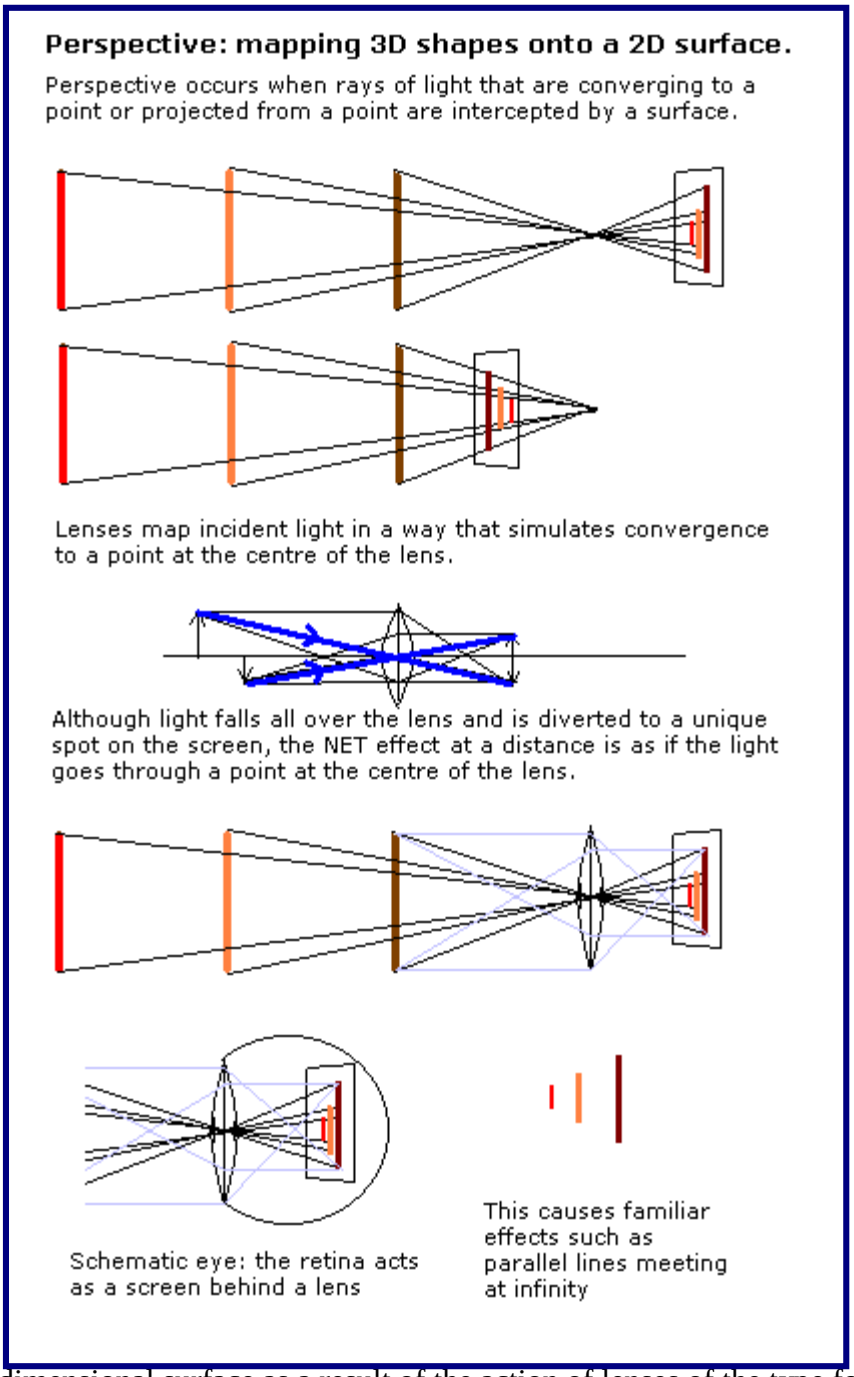
The eye is about 23 mm deep from the front of the cornea to the back of the retina. The refractive index of the components of the lens system varies from about 1.33 to 1.39.

Light from every point of a field of view falls all over the surface of the eye. There is no

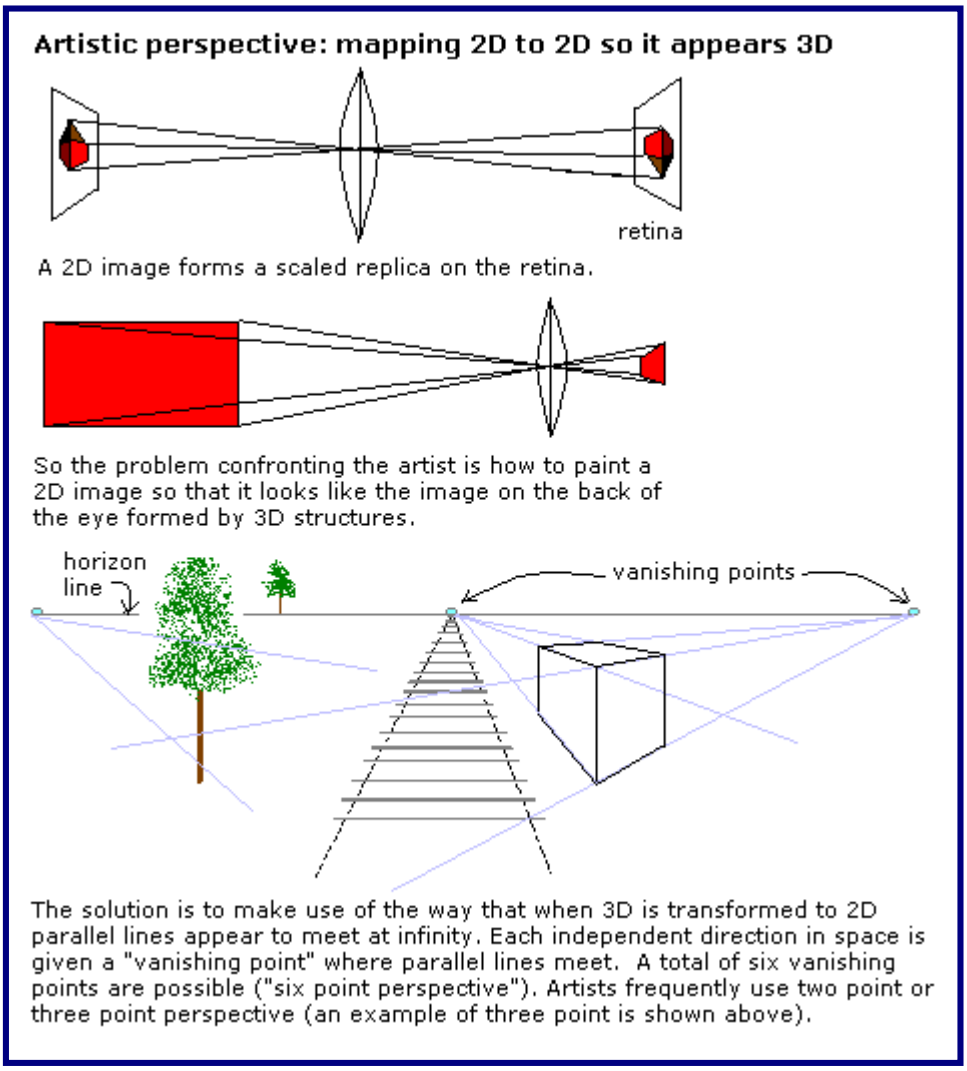
'point eye' and there is no ordered image between objects in the view and the retina except on the retina. The image on the eye has the form of an inverted mapping of 3D objects to a 2D surface. This is also the form of conscious experience so the images on the retinas are the closest physical analogues of phenomenal, visual, conscious experience (see Perspective below).

Perspective

Perspective describes how light from three dimensional objects is mapped onto a two dimensional surface as a result of the action of lenses of the type found in the eye.



Perspective is used by artists to create the impression of viewing a 3D scene. To do this they create a 2D image that is similar to the image on the retina that would be created by the 3D scene.



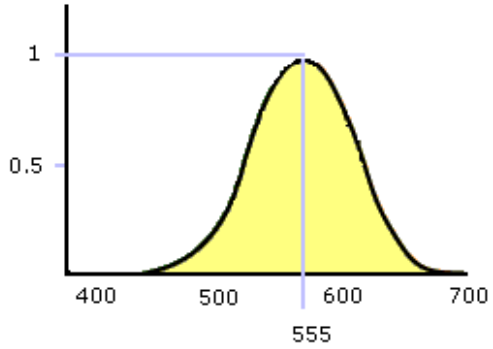
Naive Realists and many Direct Realists believe that the 2D perspective view is the way things are actually arranged in the world. Of course, things in the world differ from images because they are arranged in three dimensions.

Colour

The colour of an object can be represented by its **spectral power distribution** which is a plot of the power available at each wavelength. The unit of light power is the watt but the unit that is used to measure subjective illumination is the **candela**. One candela is the illumination due to light of a wavelength of 555 nanometres and a radiant intensity of 1/683 watts per steradian in the direction being measured. A steradian is a solid angle at the centre of sphere of one metre radius that is subtended by one square metre of the surface. The curious number 1/683 occurs because the unit was originally based on light emitted from a square centimetre of molten platinum. The wavelength of 555 nm is chosen because this is

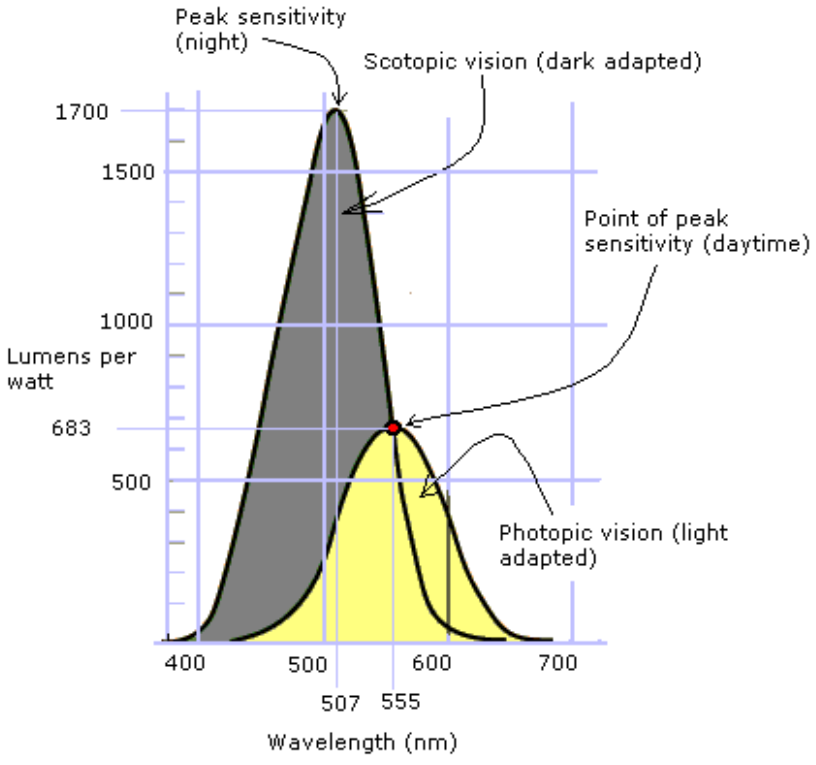
the wavelength of peak sensitivity for light adapted (photopic) vision over a large group of subjects. Light adapted vision is largely due to photosensitive cells in the retina called **cones**. The candela is fixed as a standard SI Unit for light at a wavelength of 555 nanometres. The **lumen** is a subjective measure of the flux of light energy passing through a solid angle (a steradian). 683 lumens of light at 555 nm are equivalent to a watt passing through the solid angle. At a wavelength of about 520 nm only 500 lumens of luminous flux occur per watt because the visual system is less sensitive at this wavelength. The curve of sensitivity of the visual system to light is known as the **V-lambda Curve**. At a wavelength of about 510 nm the same radiant intensity is seen as being half as bright as at a wavelength of 555 nm.

The V-lambda curve



The V-lambda curve is the curve of relative sensitivity of the visual system at different wavelengths. It is obtained from subjective comparisons of the brightness of a source at a test wavelength with the brightness of a source of the same power at the wavelength of peak sensitivity. The curve shown above is for photopic vision, another curve applies for scotopic vision.

Spectral luminous efficacy



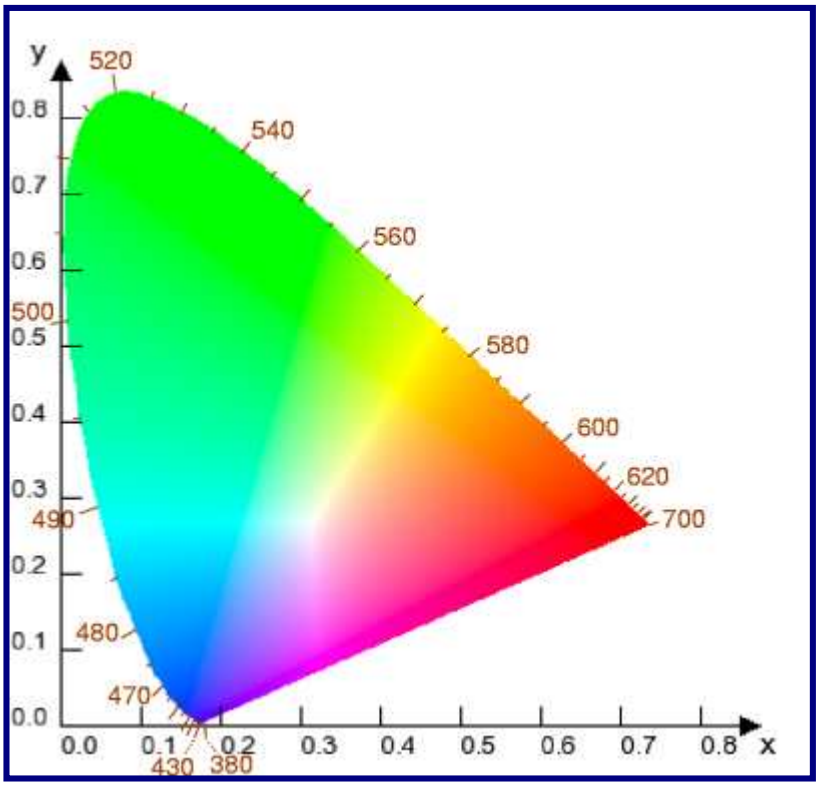
Dark adapted (scotopic) vision has a peak sensitivity at a wavelength of 507 nm and is largely due to photosensitive cells called **rods** in the retina. Spectral Luminous Efficacy Curves are also used to express how the sensitivity to light varies with wavelength.

Phenomenal colours are due to mixtures of **spectral colours** of varying intensities. A **spectral colour** corresponds to a wavelength of light found on the electromagnetic spectrum of visible light. Colours have three attributes: **brightness**, **saturation** and **hue**. The brightness of a colour depends on the illuminance and the reflectivity of the surface. The saturation depends on the amount of white present, for instance white and red make pink. The hue is similar to spectral colour but can consist of some combinations - for instance magenta is a hue but combines two spectral colours: red and blue. It should be noted that experiences that contain colour are dependent on the properties of the visual system as much as on the wavelengths of light being reflected.

Any set of three colours that can be added together to give white are known as **primary colours**. There are a large number of colours that can be combined to make white, or almost any other colour. This means that a set of surfaces that all appear white could reflect a wide range of different wavelengths of light.

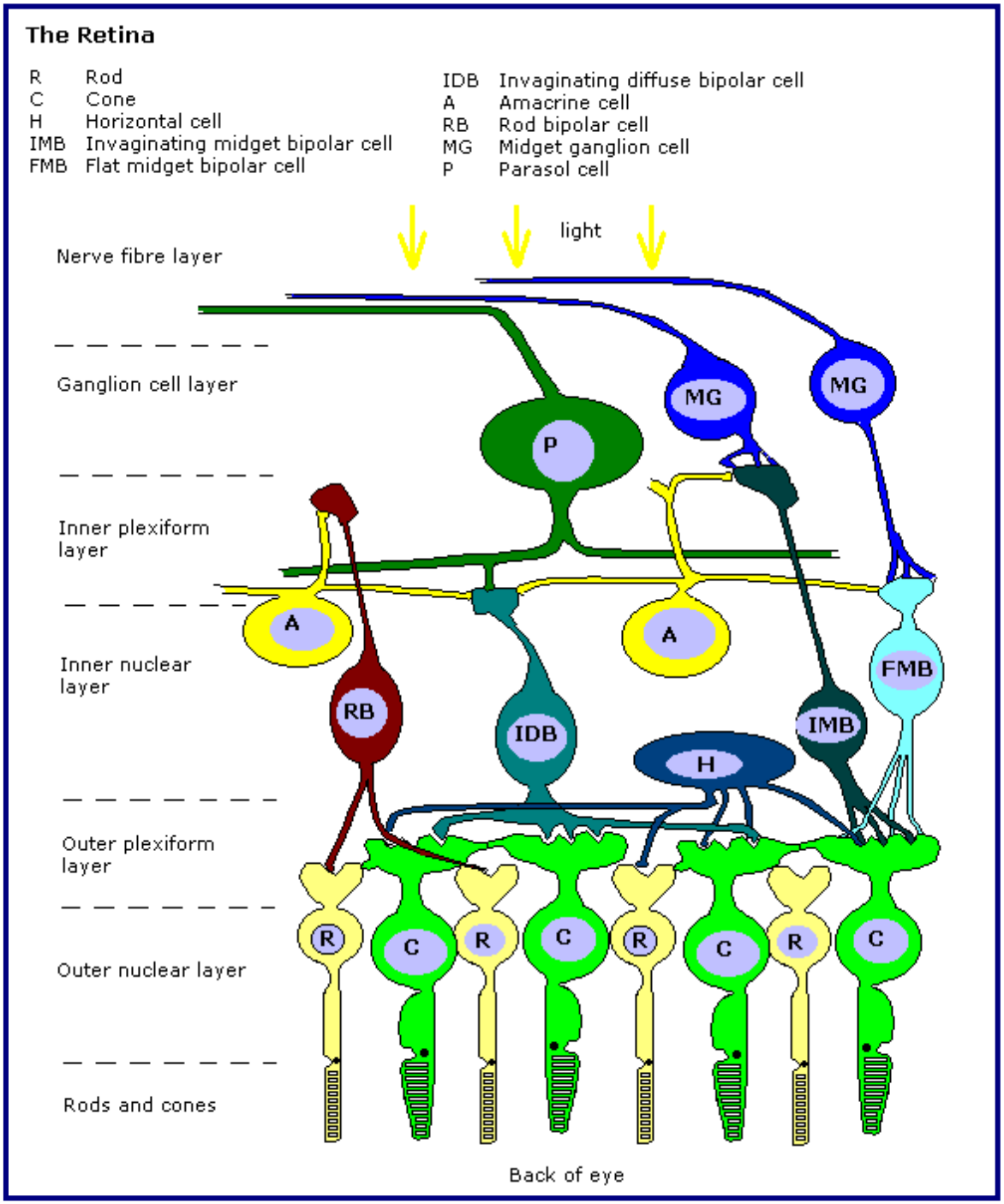
There are numerous systems for predicting how colours will combine to make other colours; the CIE Chromaticity Diagram, the Munsell Colour System and the Ostwald Colour System have all been used. The 1931 CIE Chromaticity Diagram is shown below:

See Chromaticity diagram in Wikipedia for more information.



The retina

The retina contains photoreceptive cells called rods and cones and several types of neurons. The rods are generally sensitive to light and there are three varieties of cones sensitive to long, medium and short wavelengths of light (L, M and S type cones). Some of the ganglion cells in the retina (about 2%) are also slightly light sensitive and provide input for the control of circadian rhythms. A schematic diagram of the retina is shown below.

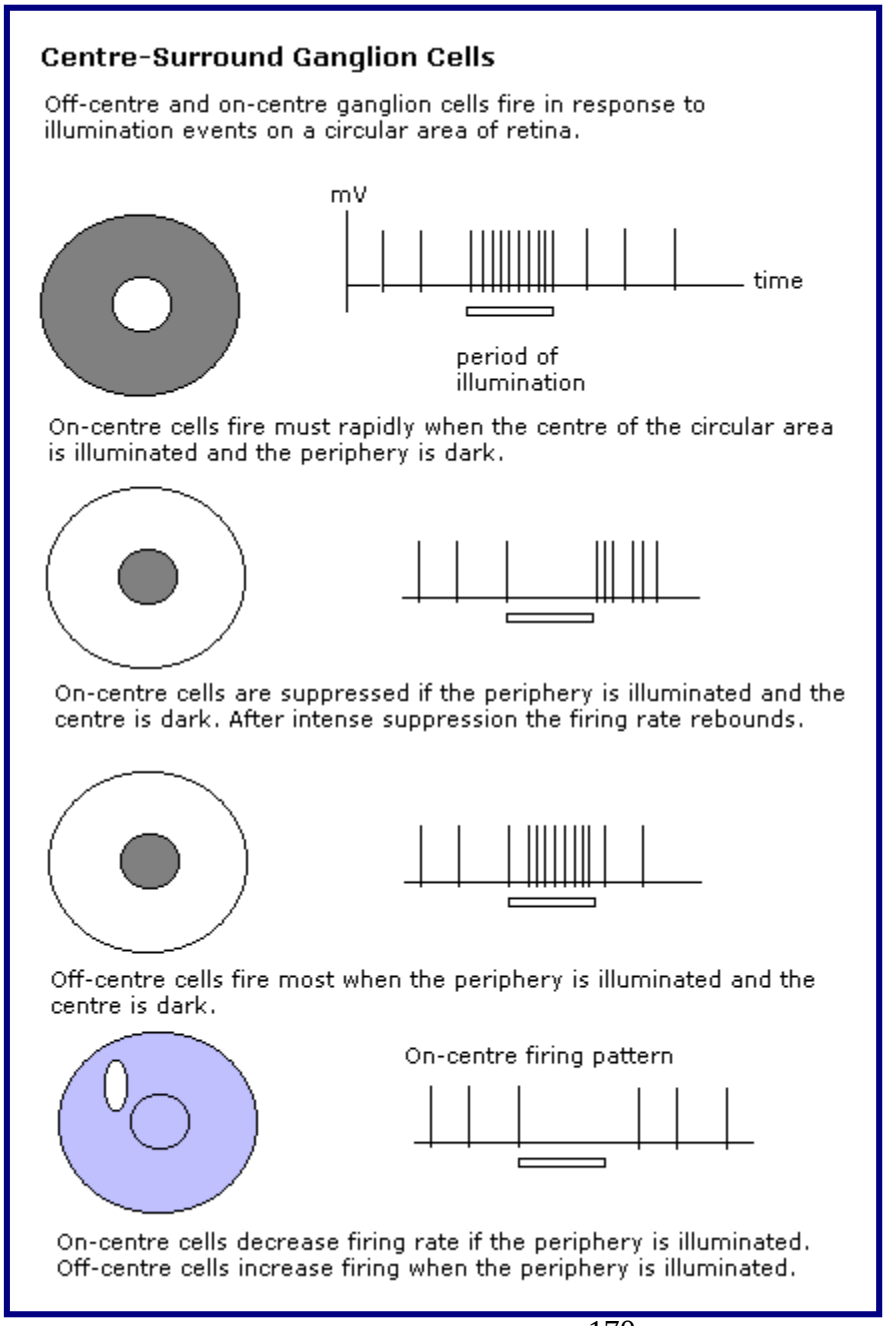


The photoreceptors hyperpolarise (their membrane potential becomes more negative) in

response to illumination. Bipolar cells make direct contact with the photoreceptors and come in two types, *on* and *off*. The on-bipolar cells are also known as *invaginating* bipolars and the off-bipolars as *flat bipolars*. On-bipolars depolarise when light falls on the photoreceptors and off-bipolars hyperpolarise. Action potentials do not occur in the bipolar or photoreceptor cells.

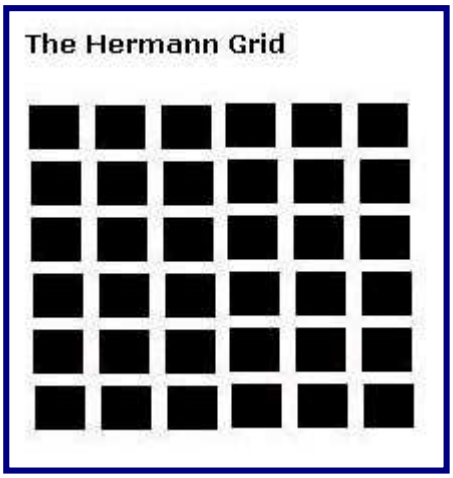
The retinal neurons perform considerable preprocessing before submitting information to the brain. The network of horizontal and ganglion neurons act to produce an output of action potentials that is sensitive to boundaries between areas of differing illumination (edge detection) and to motion.

Kuffler in 1953 discovered that many retinal ganglion cells are responsive to differences in illumination on the retina. This **centre-surround** processing is shown in the illustration below.

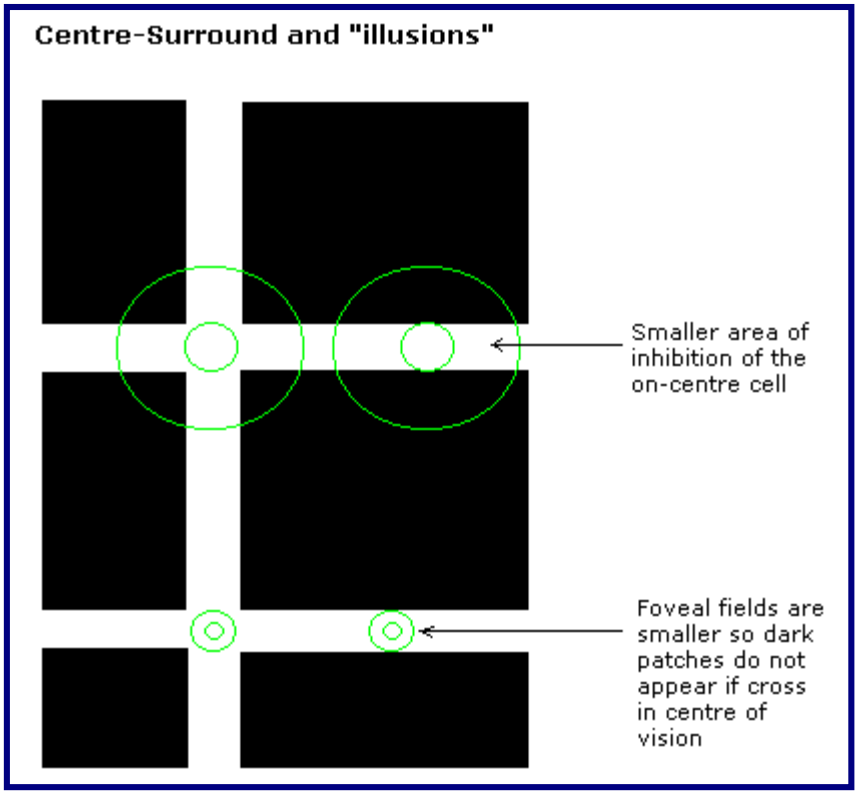


The centre-surround effect is due to **lateral inhibition** by horizontally arranged cells in the retina.

The structure of the response fields of ganglion cells is important in everyday processing and increases the definition of boundaries in the visual field. Sometimes it gives rise to effects that are not directly related to the physical content of the visual field. The most famous of these effects is the Hermann Illusion. The Hermann Grid Illusion is a set of black squares separated by white lines. Where the white lines cross it appears as if there are grey dots.

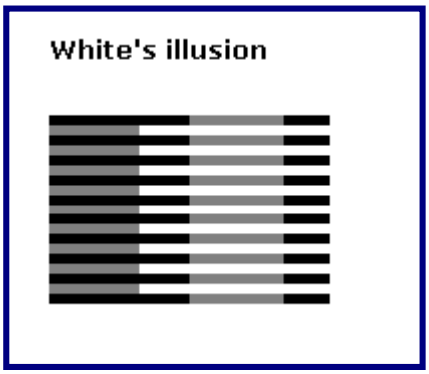


The grey dots are due to the relative suppression of on-centre ganglion cells where the white lines cross. This is explained in the illustration below.



Notice how the grey dots disappear when the crossed white lines are at the centre of the visual field. This is due to way that ganglion cell fields are much smaller in the fovea.

There are many other retinal illusions. White's illusion is particularly strong and also due to centre-surround activity.



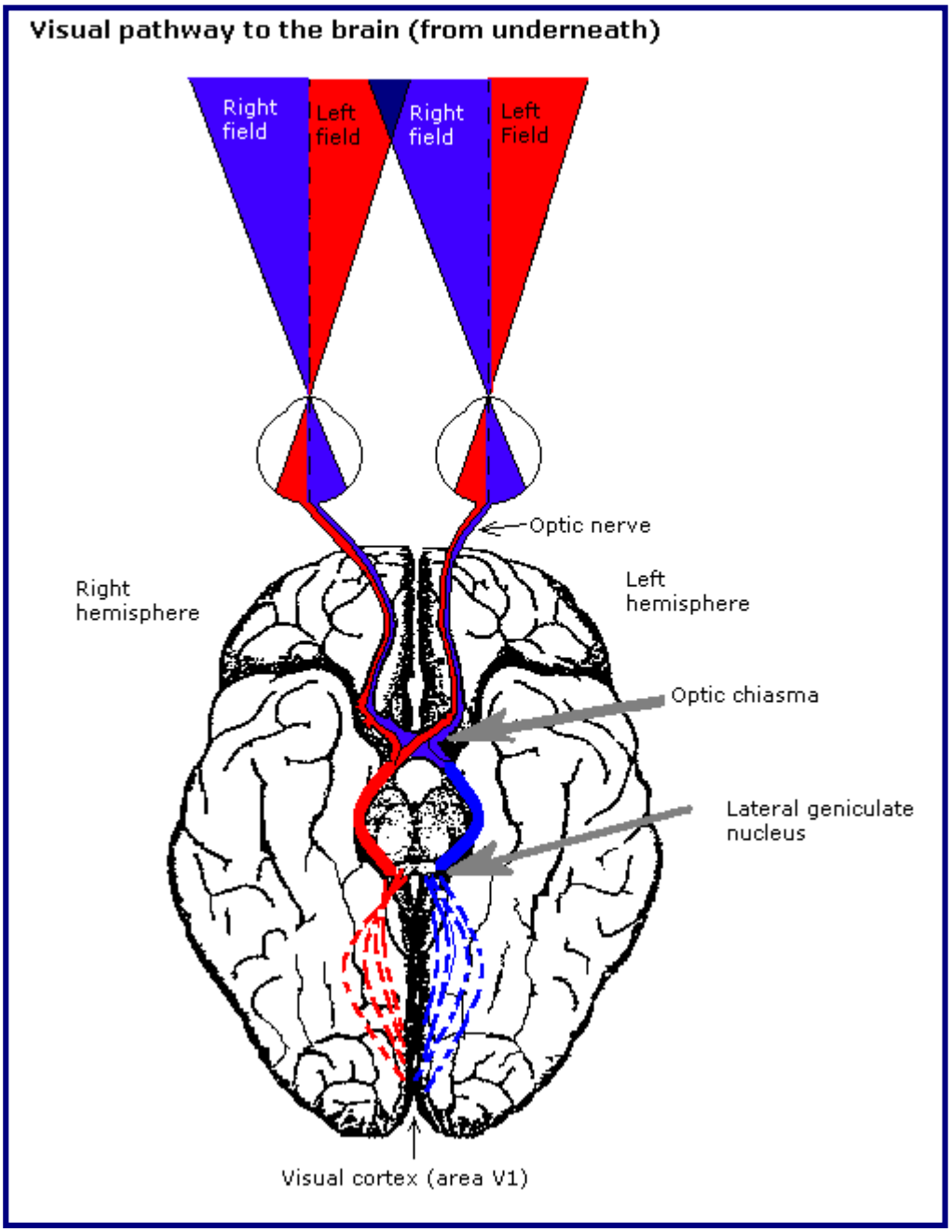
The grey lines really are the same shade of grey in the illustration. Mach's Illusion is another example of a centre-surround effect. Centre-surround effects can also occur with colour fields, red/green and yellow/blue contrasts having a similar effect to light/dark contrasts.

Lateral inhibition and the resultant centre-surround effect increases the number of cells that respond to boundaries and edges in the visual field. If it did not occur then small boundaries might be missed entirely if these fell on areas of the retina outside of the fovea. The result of this effect is everywhere in our normal visual phenomenal experience so not only is visual experience a mapping of 3D on to a 2D surface, it also contains shading and brightening at edges that will not be found by photometers that measure objective light intensities.

Photoreceptors become less responsive after continuous exposure to bright light. This gives rise to **afterimages**. Afterimages are usually of the opponent colour (white light gives a dark afterimage, yellow light gives a blue afterimage, red gives a green afterimage etc). Afterimages when the eyes are open are generally due to a lack of response to a particular frequency of light within the white light that bathes the retina.

It is clear that visual phenomenal experience is related more directly to the layout and type of activity in the retinal cells than to things in the visual field beyond the eye.

Visual pathways



The lateral geniculate nucleus

Retinal ganglion cells project to the Lateral Geniculate Nuclei which are small bumps on the back of the thalamus. (Only 10-15% of the input to the LGN comes from the retina, most (c.80%) comes from the visual cortex). The neurons in the LGN are arranged retinotopically so preserve the layout of events on the surface of the retina.

The LGN are arranged in 6 layers. The top two are known as Magnocellular layers (about 100,000 neurons with large cell bodies) and the bottom four are called Parvocellular layers (about 1,000,000 neurons with small cell bodies). Between the main layers are the Koniocellular layers that consist of large numbers of tiny neurons.

The left Lateral Geniculate Nucleus receives input from the right visual field and the right LGN receives input from the left visual field. Each nucleus receives input from both eyes but this input is segregated so that input from the eye on the same side goes to layers 1, 3, 5 and from the other side to layers 2,4, 6.

The magnocellular layers contain neurons that have a large receptive field, are sensitive to contrast, a transient reponse and are not colour sensitive. The parvocellular layers contains neurons that have small receptive fields, are colour sensitive, have a prolonged response and are less sensitive to contrast.

The LGN pathway from the retina is largely connected to the striate part of the visual cortex (cortical area V1) via a set of fibres called the optic radiation. There are reciprocal connections between the Thalamic Reticular Nucleus and the LGN. The LGN are also interconnected with the Superior Colliculus and brainstem.

The LGN may be involved in controlling which areas of the visual field are subjected to attention (O'Connor et al 2002).

The visual cortex

The input from the LGN goes mainly to area V1 of the cortex. The cortex is arranged in six layers and divided up into **columns**. Each column in the visual cortex corresponds to a particular area of the retina in one eye. The columns are arranged in rows called **hypercolumns**. Each column within a hypercolumn responds to a different orientation of an optical stimulus at a given location (so responds to edges/boundaries that are oriented in the visual field). Hypercolumns from each eye are arranged alternately and form a small block of cortex called a **pinwheel**. At the centre of each pinwheel are colour sensitive cells that are usually not orientation sensitive. These coincide with the "blobs" that are seen when visual cortex is viewed using cytochrome oxidase dependent stains. It is important to note that the "hypercolumns" merge into one another and respond to line stimuli that cover an area of retina so they may be physiological rather than anatomical entities.

The blind spot in each eye is represented by an area of visual cortex that only receives monocular input from the other eye (Tong & Engel 2001). The effect of the blind spot is illustrated below:

"Filling in" of the blind spot

Concentrate on the cross with the left eye whilst the right eye is closed, at some distances and tilts of the head the black circle disappears (similarly focussing on the black circle with the right eye can make the cross disappear).

When both eyes are used the cortical area representing the blind spot for the eye with the missing data is filled in with data from the other eye. When one eye is used the phenomenal experience contains the colour of the background.

Normally it seems that the blindspot is 'filled in' with background when one eye is used. However, Lou & Chen (2003) demonstrated that subjects could respond to quite complex figures in the blind spot, although how far they were investigating 'blindsight' rather than visual experience in the blind spot is difficult to determine.

Different layers in the visual cortex have outputs that go to different locations. Layer 6 sends nerve fibres to the Lateral Geniculate Nuclei and thalamus, layer 5 to superior colliculus and pons, layer 2 & 3 to other cortical areas.

There are two important outputs to other cortical areas, the **ventral stream** and the **dorsal stream**. The ventral stream processes colour, form and objects. It proceeds to the inferior (lower) temporal cortex. The dorsal stream processes motion, position and spatial relationships. It proceeds towards the parietal cortex. Lesions in the ventral stream can result in patients knowing where an object is located but being unable to enumerate its properties, on the other hand, lesions to the dorsal stream can result in patients being able to label an object but unable to tell exactly where it is located.

There is also a large output from the visual cortex back to the thalamus, this output contains more fibres than the thalamo-cortical input.

Depth perception

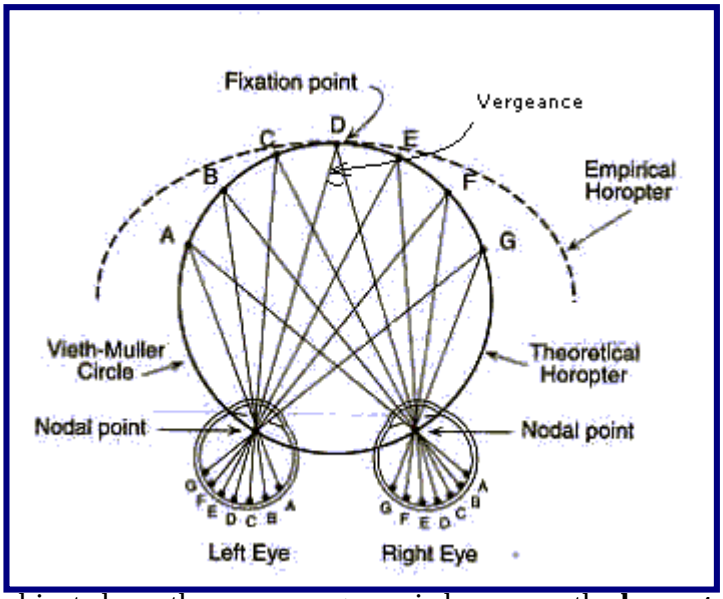
The world is three dimensional but the image on the back of the retinas is two dimensional. How does the brain give the subject a perception of depth?

Depth perception relies on **cues** which are data about the displacement of things relative to the body. These cues consist of:

- the convergence of the eyes

- the accommodation of the lens
- binocular disparity -the difference between the images on the retinas- this was first suggested by Wheatstone.
- motion parallax - distant objects move slower when the observer moves - first suggested by Helmholtz.
- optical flow - the rate of expansion/contraction of a scene with movement towards or away from it (Lee & Aronson 1974).
- binocular occlusion - parts of a scene are invisible to each eye.
- body motion provides cues about near objects.
- vanishing points - the convergence of parallel lines.
- numerous other cues such as size constancy, texture etc.

Binocular disparity has been most extensively studied as a source of depth cues. When the eyes converge to focus on an object in from of them there is very little disparity in the images of that object on the two retinas. The angle at the object formed between the lines that project back to the pupils is known as the **vergence** at the object. The sphere where all



objects have the same vergence is known as the **horopter**.

When the disparity between the retinas is small a single image occurs in phenomenal experience which is accompanied by a sensation of objects with depth. This is known as **stereopsis**. If the disparity between the retinas is large double vision ensues, this is known as **diplopia**. The curious feature of stereopsis is that we can see no more of the object than is visible on the retinas and certainly cannot see behind the object. Stereopsis is more like a stretching of 2D space than actual 3D.

In the review by Cutting and Vishton (1995) the contributions of each type of cue is discussed. Cutting and Vishton also present evidence that there are several zones of depth perception that are informed by different sets of cues. These are **personal space**, which is the zone of things within arms reach, **action space**, which is the zone in which we interact and where our motions have a large impact on the perceived layout, and **vista space** which is the zone beyond about 30m that is informed by long range cues.

The interesting feature of perceptual space is that it is not seen. The sides of a solid object appear as intrusions or lateral extensions in 2D space, when we close an eye that has access to the side of the object and then open it again the side grows out into 2D space. The lack of 'seeing' depth is also evident when we close one eye when looking at a vista - nothing seems to change even though stereopsis has gone. This leaves the problem of what it is that constitutes the 'feeling' of depth. We have feelings that we can fall into space or move into it or around in it. Depth seems to be defined by premotor modelling and the potential for occupancy by our bodies and limbs. As such it involves qualia that are different from those of vision and more akin to those that accompany movement, as an example, if you reach out to touch something, move the hand back, then consider the distance to the object it is evident that a feeling of the movement is still present. Is depth a quale of movement modelled during the extended present of perception?

References

- Cutting, J.E. & Vishton, P.M. (1995) Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein & S. Rogers (eds.) Handbook of perception and cognition, Vol 5; Perception of space and motion. (pp. 69-117). San Diego, CA: Academic Press. http://pmvish.people.wm.edu/cutting&vishton1995.pdf

- Gregory, R.L. (1997). Knowledge in perception and illusion. From: Phil. Trans. R. Soc. Lond. B (1997) 352, 1121â€“1128. http://www.richardgregory.org/papers/knowl_illusion/knowledge-in-perception.pdf

- Hubel, D.H. (1981). EVOLUTION OF IDEAS ON THE PRIMARY VISUAL CORTEX, 1955-1978: A BIASED HISTORICAL ACCOUNT. http://nobelprize.org/medicine/laureates/1981/hubel-lecture.pdf

- Lou, L. & Chen, J. (2003). Attention and Blind-Spot Phenomenology. PSYCHE, 9(02), January 2003. http://psyche.cs.monash.edu.au/v9/psyche-9-02-lou.html

- O'Connor, D.H., Fukui, M.M., Pinsk, M.A. & Kastner, S. (2002). Attention modulates responses in the human lateral geniculate nucleus. Nature Neuroscience 5, 1203 - 1209 (2002). http://www.nature.com/neuro/journal/v5/n11/full/nn957.html#B34

- Tong, F., & Engel, S. A. (2001). Interocular rivalry revealed in the human cortical blind-spot representation. Nature, 411, 195-199. http://www.psy.vanderbilt.edu/tonglab/publications/Tong&Engel2001.pdf

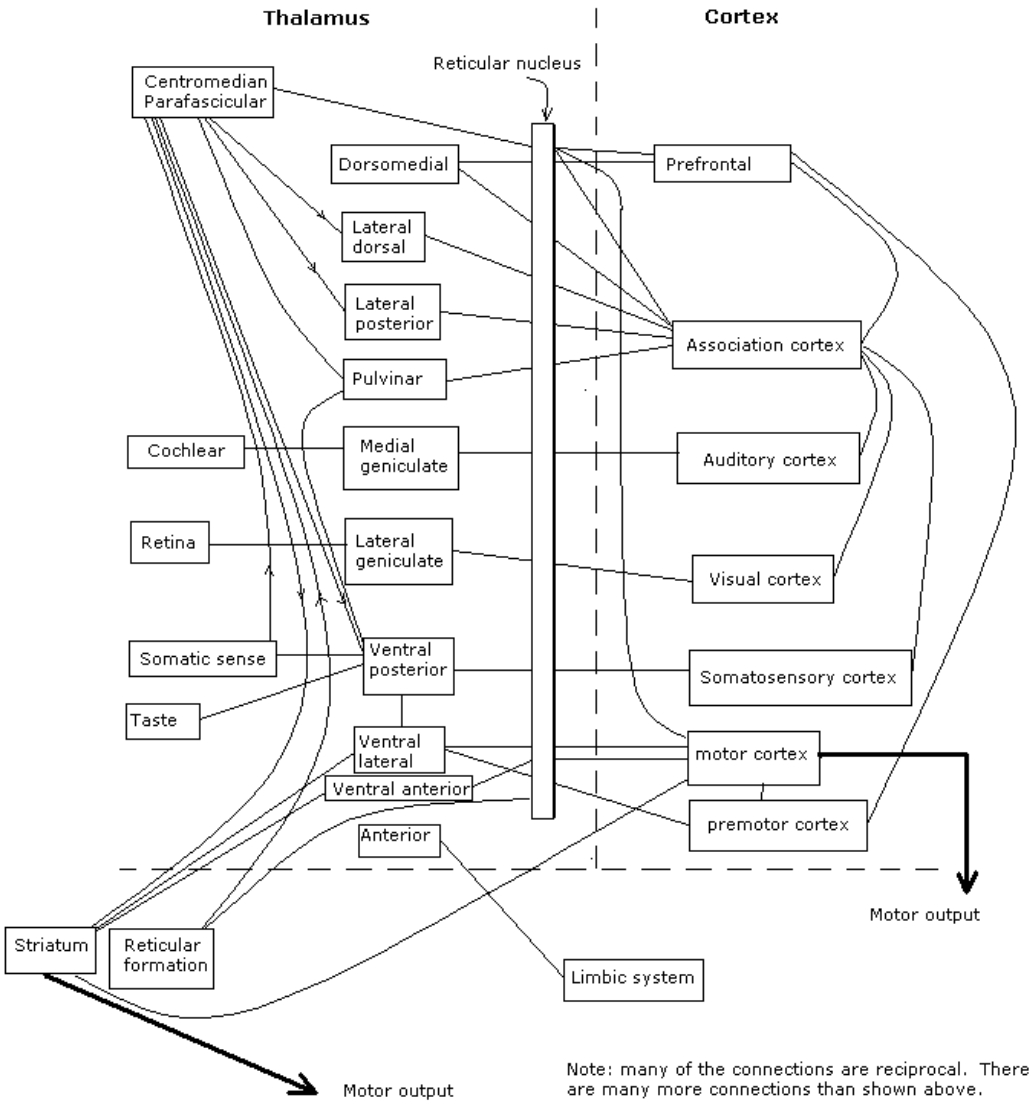
- Tong, F. (2003). PRIMARY VISUAL CORTEX AND VISUAL AWARENESS. Nature Reviews of Neuroscience. 4:219-229. http://www.psy.vanderbilt.edu/tonglab/publications/Tong_NRN2003.pdf

The Cortex and Thalamus

The cortex and consciousness

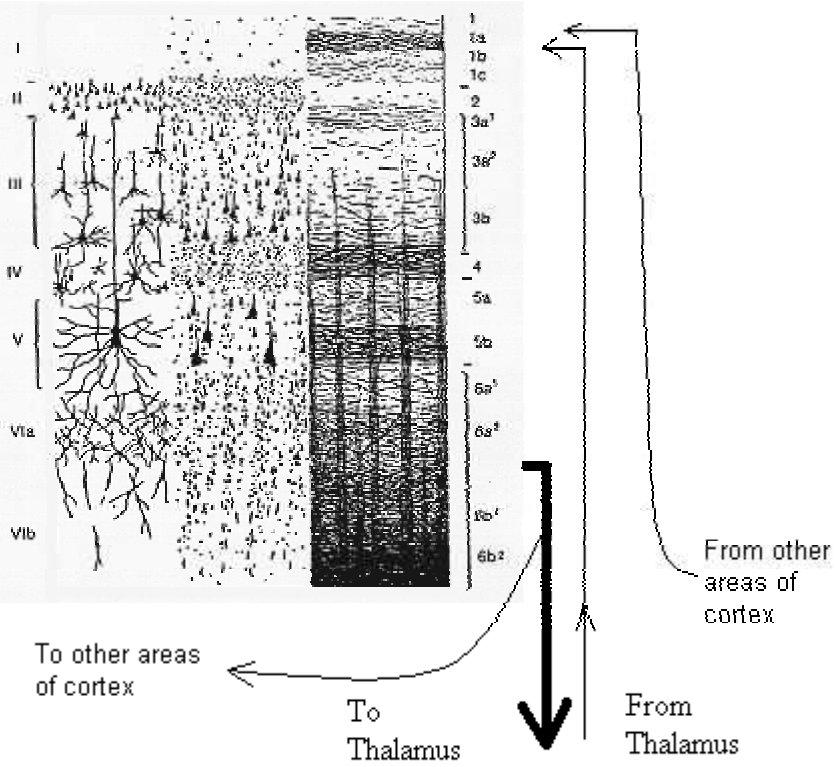
The cerebral cortex consists of a set of specialised areas that process different aspects of sensation and motor control. There are about ten times as many nerve fibres going from the cortex to the thalamus as there are from the thalamus to the cortex (Destexhe 2000).

The sensory processors of the diencephalon and telencephalon showing the two synergistic processor networks.



Histologically the cerebral cortex is a layer of greyish neurons overlying a huge mass of white nerve fibres, the cerebral medulla. The cortex consists of six main layers. The upper layers receive input from the relays in the thalamus such as the lateral geniculate, from the

thalamus in general and from other areas of cortex plus a few specialised inputs from other locations. The lower layers give rise to output fibres that largely connect with the thalamus and other areas of cortex although particular specialised processors in the cortex may also have direct connections elsewhere such as to motor nuclei.



The cerebral cortex has many functions and is divided up into numerous separate processors. The most important function of the cortex from the point of view of consciousness studies is that it creates models. These models are most easily experienced when there is a lack of sensory input such as in dreaming, day dreaming, lucid dreaming or experiencing imaginary speech (thinking). In ordinary waking life the modelling processes create a model of the world around us and within us based on sense data and associated data. This model consists of overlapping sounds, images, smells etc. and is a combination of perceptual fields from all the senses.

There is considerable evidence that the parts of the brain that deal with imagining (modelling) things are also the parts that deal with perception (ie: modelling the world). The overlap between imagination and normal perception is not complete because, as Tong(2003), in a review of visual consciousness, put it: "Internally generated experiences share some, but not all, of the phenomenal properties of actual perception". There is also considerable overlap between the areas used for imaginary speech (thought) and actual speech, areas dealing with the control of sensation and of the tongue etc. being used in actual speech but not in imagined speech (Fu et al 2002). Kreiman et al (2000) investigated the activity of single neurons in humans and also found that the brain activity evoked by visual imagination overlapped that which occurs upon direct stimulation by the same image.

Our conscious experience consists of the output of the cortical modelling processes. The cerebral cortex itself appears to be non-conscious. The evidence for the non-conscious nature of the cerebral cortex is reinforced by lesion studies that show that up to 60% of the cerebral cortex can be removed without abolishing consciousness. Either hemisphere can be removed or much of the front or back of the cerebral cortex can be cut off yet consciousness persists. The cerebral cortex is often assumed to be the "seat" of consciousness because this collection of organs is relatively large in humans but the truth seems to be that the cortex is a collection of processors that provide an input to experience. There is also a substantial amount of neurophysiological evidence that the cortex is non-conscious.

Libet et al (1967) found that there could be cerebral cortical activity in response to weak stimulation of the skin without any conscious awareness of the stimulus. This work provides a neurophysiological basis for subliminal (non-conscious) perception and also shows that large areas of the cerebral cortex can be active without conscious experience. The insensitivity of experience to cortical activity has been further confirmed by Libet et al (1979). They electrically stimulated the cerebral cortex of conscious patients and discovered that the stimulus must be continued for about 0.5 seconds for subjects to report a conscious experience of the stimulation. What is the cortex doing in the 0.5 seconds between the start of stimulation and the report of awareness of the stimulation? It is probably synchronising its various processors and creating a waking dream, a structured set of events that accounts for the activity.

The 'Attentional Blink' (Raymond et al 1992) is also consistent with the concept of the cerebral cortex being a device that creates models. In the 'Attentional Blink' the identification of an object impairs the identification of a second object that is presented within 0.5 seconds of the first. Raymond et al used a stream of letters (11 letters per second) and the identification of a first letter impaired the identification of a subsequent 'probe' letter in the stream. If the probe letter followed the first letter within about 180 msecs it could easily be identified, suggesting that chunks of about 180 msecs of data stream are modelled together. Christmann & Leuthold (2004) have theorised that the 'Attentional Blink' involves perceptual and central components of visual processing. This is supported by the fMRI studies of Marois et al (2004) who presented subjects with faces mounted on scenes of places. The scenes of places often went undetected by subjects but they activated regions of the medial temporal cortex involved in high-level scene representations, the parahippocampal place area (PPA). When the scenes of places were detected by the subjects there was activity in the frontal cortex and the PPA activity was increased. These experiments are consistent with the idea of a cerebral cortex that is a multiprocessor system that creates consistent models of the environment for presentation to some other part of the brain.

Bregman's (1990) auditory continuity illusion is another example of how sensory events are modelled. If a pure tone is followed by broadband noise and the noise followed by the same pure tone it seems as if the tone occurs throughout the period of noise. If the noise is not followed by the pure tone there is no sound of the tone during the period of noise. This effect is similar to the results found by Libet because a delay of several hundred milliseconds between sensory stimulation and conscious experience is needed to account for the apparent rewriting of history after the second tone appears.

The 0.5 second delay required for the cortex to model an event has implications for the role

of conscious experience in the control of our lives. If experience is always 0.5 seconds behind the true present instant then how can we be said to control anything? The brain must be acting automatically whilst performing most tasks. The 0.5 second delay also seems to contradict our everyday experience. We certainly feel like we are aware of things in less than 0.5 seconds, for example, the direct stimulation of sense organs seems to be experienced much more rapidly than the delayed experience of cortical stimulation. In fact subjects report that they are conscious of stimuli, such as being touched or seeing flashing lights, within 0.1 to 0.2 seconds of the event. So how can subjects report events within 0.2 seconds even though it seems to take 0.5 seconds for the cortex to generate activity that can be experienced? The simplest explanation is that the reaction occurs automatically within 0.2 seconds and then the conscious experience of this reaction occurs 0.3 seconds later. This gives a total 0.5 seconds delay before conscious experience whilst allowing fast reactions.

Libet et al extended their experiments by stimulating a "relay nucleus" in the thalamus that intercepts signals from the senses before they reach the somatosensory cortex. It was found that when this nucleus was stimulated for 0.5 seconds the subjects reported that the stimulus occurred 0.2 seconds after it had begun. When the nucleus was stimulated for less than 0.5 seconds the subjects did not report any sensation. This supports the concept of a 0.5 second delay whilst the cortex puts a stimulus in context before it is experienced.

These experiments show that our experience is an output of cortical processing rather than the processing itself. If our conscious experience is non-cortical then this raises the possibility that the non-conscious cerebral cortex can perform actions without conscious control. Of course, the cortex does this all the time when we are indulging in skilled or routine behaviour. The ability of the non-conscious cortex is quite remarkable; for instance car drivers sometimes discover that they have driven for several miles without conscious experience of driving, even at the level of having no recollection of the route.

Although it might be accepted that much of our everyday behaviour is automatic is there any behaviour that is definitely initiated by conscious experience? This is probably a pointless question because consciousness is about observation, not action; however, despite this there have been several experiments that have attempted to determine the relationship between consciousness and action.

In 1964 Kornhuber and Deecke performed a series of experiments that measured the electrical activity from the scalp (EEG) during voluntary actions. They averaged many EEG's from subjects who were about to move a finger and discovered that there is an increase in scalp potential before the movement takes place. The increase in potential can start as long as 2 seconds or so before the movement and is known as the "readiness potential" (Bereitschaftspotential). The readiness potential is strange because it seems to contradict our conscious experience; we do not decide to move a hand and then wait 2 seconds before the hand moves. It seems that the non-conscious brain may be taking things into its own hands.

Libet et al (1983) extended the readiness potential experiments by asking subjects to observe a Wundt clock whilst flexing a finger. The Wundt clock had a spot of light that moved around a circle every 2.56 seconds and allowed the subjects to obtain timings that were related to their mental experiences. When the subjects flexed a finger it was found that the readiness potential occurred about 0.5 seconds before the finger moved and the subjects reported they were going to move the finger about 0.2 seconds before the movement. This

suggested that a subject's cerebral cortex was preparing for the movement about 0.3 seconds before the subject was conscious of this. Libet's experiments have been reproduced elsewhere (see Keller & Heckhausen 1990). (It is important to note that the subjects in Libet's experiment were asked to wait until they felt the urge to move the finger.) These results are consistent with the idea of the cortex as a modelling system that constructs a consistent model of events to pass on to whatever mediates conscious experience.

Perception, Imagination, Memory and Dreams

Functional Magnetic Resonance Imaging (fMRI) has shown that similar areas of brain are used during perception involving the senses as during imagination (Tong 2003, Kosslyn and Thompson 2003). The phenomenal substrate of the mental images that occur in both modes of brain activity has not yet been found.

Ganis et al (2004) used fairly complex perceptual and imagination tasks that activated large areas of the brain, they found an overlap between the brain areas activated during perception and imagery. The principle areas that were different in the two tasks were found in the primary sensory areas of the visual cortex. Other areas in the visual cortex and activity in the rest of the brain showed a remarkable degree of overlap. The authors suggested that the differences in the activity of primary visual cortex may have been due to differences between the perceptual and imaginary stimuli such as speed of onset etc. The hippocampus was not activated.

It is intriguing that, contrary to object imagery, spatial imagery such as predicting when a cross on a screen would fall on an imaginary letter actually seems to inhibit activity in sensory visual cortex (Aleman et al). Both fMRI and blocking with transcranial magnetic stimulation (TMS) showed that the posterior parietal cortex was involved in the spatial imagery.

Imagery involving places and faces activates the place and face areas that are activated during perception (Ishai et al 2000).

The recall and recognition of things also seems to involve very similar brain areas to those used during perception. Wheeler and Buckner (2003) showed that areas involved in perception were also involved in the recall of the perceptual stimuli.

Recall causes activation of areas used in perception but also seems to use areas that may be particularly related to the process of recall iself, such as the left parietal cortex (Konishi et al 2000) (Brodmann's area 40/39). Frontal and parietal regions are involved in the recognition of whether stimuli have been experienced before.

Image generation during sleep seems to differ from that during imagination and recall. In particular it seems to involve a few well defined areas of cortex and considerable activation of the posterior thalamus.

Sleep studies have shown that people dream throughout sleep. However, dreams are more frequent during the REM (rapid eye movement) periods of sleep than the NREM (non-REM) periods. Dreams are reported after 70-95% of awakenings in REM sleep and 5-10% of awakenings in NREM sleep. REM dreams are more visual than NREM dreams which are more 'thoughtlike' (Solms 2000). Thoughtlike events (mentation) are reported after 43% of

awakenings from NREM sleep.

Solms (1997) found that patients who had lesions in the parietal-temporo-occipital junction reported a cessation of visual images in dreams. Solms also found that patients with lesions in the white matter inferior to the frontal horns of the lateral ventricles, in the ventromesial quadrant of the frontal lobes, also reported loss of dreaming. Loss of dreaming is also reported by leucotomised patients with frontal ventromesial damage. Damasio et. al. (1985) and Solms (1997) also reported that some patients with damage to the medial prefrontal cortex, the anterior cingulate cortex, and the basal forebrain became confused about what was real life and what was dreaming (waking dreams occurred).

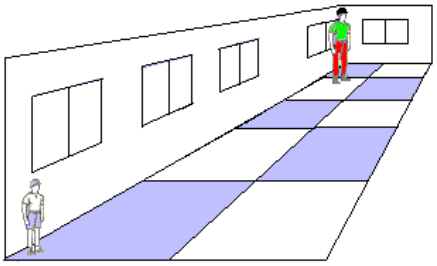
Studies using fMRI show that the sensory occipital lobe (BA 18) and posterior thalamus, especially the lateral geniculate nuclei, are activated in REM sleep, weaker activations of the posterior cingulate, putamen and midbrain were also found (Wehrle et al 2005, Loveblad et al 1999). These findings are consistent with activation of the ponto-geniculo-occipital system (PGO) during REM.

So dreams may be more like primary activations of sensory cortex than imagining or recall. This suggests that dreams have a thalamic origin or are managed via connections from the cortex through the thalamus to the visual cortex.

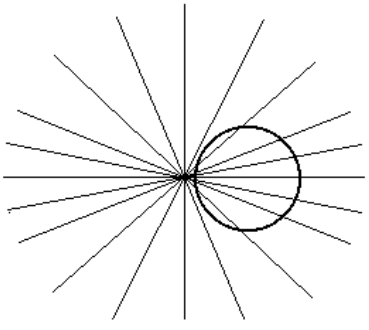
Hallucinations seem to differ from dreams. In Charles Bonnett Syndrome patients can have clear hallucinations. These, like imaginations, seem to involve areas of the visual cortex that deal with processed data, for instance hallucinations of faces activate the "face area" rather than visual cortical area V1 (ffytche et al 1998).

More about Models

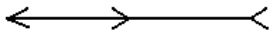
Our dreams are clearly models that form a 'dreamworld' but the idea that perception might be like a dream that is updated by sensation is not so obvious. Experience seems to be an active model of the world (virtual reality) based on sense data rather than a simple mapping of retinal and other sensory data. This is demonstrated by visual illusions such as the Ames Room, Spoke Illusion and Muller Lyer illusions shown below:



The little boy and the man are the same size on the page



The circle is symmetrical

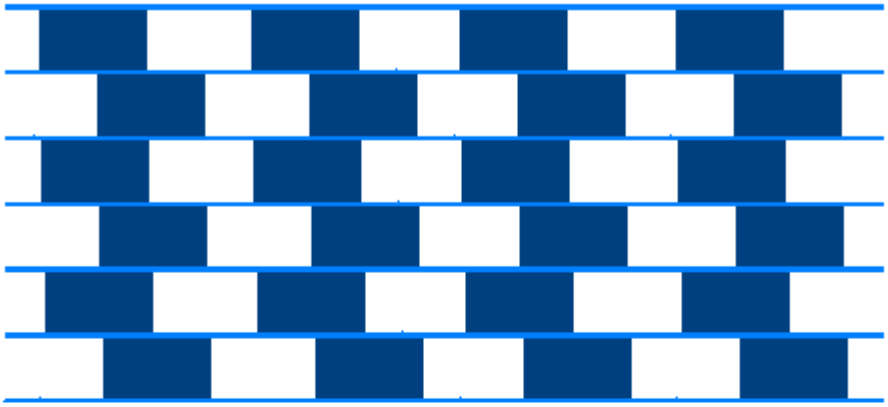


Muller Lyer: The two horizontal lines are the same length

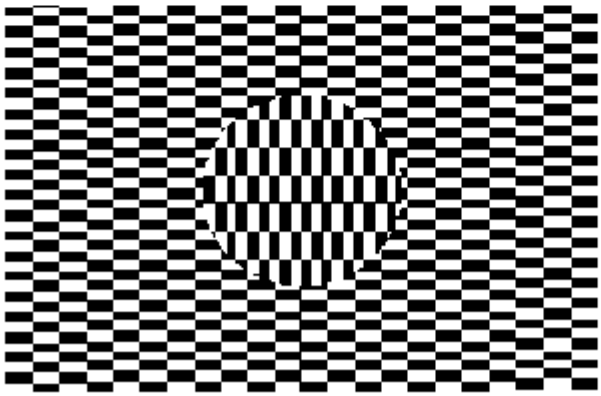
Notice how the circle is distorted without any distortion in the 'spokes', it is as if the circle

has been treated as a separate object by the processes in the brain that rearranged it. In all of these illusions the brain has rearranged large areas of the visual field and has managed the input as a collection of 'objects' that are manipulated separately. Even movement seems to occur in some figures showing that the brain models the position of things:

Cafe Wall Illusion



Ouchi's pattern



Stare at the central disk - does it move?

The creation of a model is also demonstrated by the illusion of movement experienced when we watch the cinema or television. This is due to the cortical modelling that is known as 'short-range apparent motion' rather than flicker fusion or persistence of vision. It is intriguing that, although it has been known for decades that the joining together of static images in our minds is due to modelling activity in the brain the myth that it is due to persistence of vision or flicker fusion is universal. As Anderson and Anderson (1993) noted:

> Indeed, in the past decade, psychoanalytic-Marxist film scholars have retained the model implied by persistence of vision: theirs is a passive viewer, a spectator who is "positioned," unwittingly "sutured" into the text, and victimized by excess ideology.

Our experience of the cinema is like a dream updated by sensation rather than sensation updated by interpretation. In fact the most compelling evidence for the modelling power of the brain is the existence of dreams; our dreams are often models of worlds that do not exist and involve little or no sensory input yet can involve effects as powerful as any television drama.

Short range apparent motion occurs when the interval between presentations of an object is brief (c. 50-100 msecs). Motion modelling in response to longer intervals is known as long range apparent motion. There is evidence that the modelling in short range apparent motion is enhanced if the moving patterns are similar to moving human forms (such as patterns of dots outlining a person)(Thornton et al 1998). The accuracy of predicting movement can actually improve if the interval between presentations is increased when human forms are used.

Motion modelling can also be seen in visual illusions such as the *Waterfall Illusion* (motion aftereffect). The waterfall illusion is commonly seen after viewing a sequence of scrolling credits on the television; when the credits stop rolling it appears as if they briefly move in the opposite direction. Tootel et al (1995) have used fMRI to show that this is correlated with activity in the motion modelling area of visual cortex (area MT/V5). The waterfall illusion is also associated with an intriguing aftereffect known as **storage of the motion aftereffect**. Normal motion aftereffects last for up to about ten seconds after the stimulus, however, if the subjects close their eyes for the normal duration of the aftereffect then reopen them they see the illusion for almost the normal duration. Culham et al (1999) used fMRI to show that activity in area MT/V5 was low during the period when the eyes were closed then increased dramatically when the eyes were opened. This is strongly suggestive of a modelling mechanism outside MT/V5 that has adapted to motion and then models stationary data with movement in the wrong direction.

Visual area MT/V5 is also involved in the separation of moving visual scenes into *sprites* or objects that move together as a whole within a scene (Muckli et al 2002).

The way that mental models may be the basis of ordinary reasoning was outlined by Johnson-Laird (1980), based on earlier work by Kenneth Craik.

Studies of 'change blindness' and 'inattentional blindness', where subjects fail to spot outrageous changes in their environment, also demonstrate that we experience a model and suggest that the brain must analyse an object to incorporate it fully into the model (See for instance Rensink (2000), Simons & Rensink (2005)).

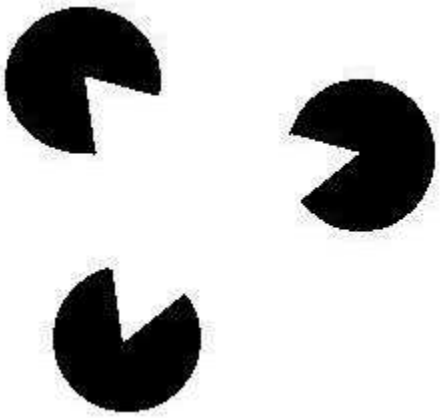
Blindsight

Blindsight studies illuminate the relationship between the cerebral cortex and our experience. When the visual cortex is removed subjects become almost totally blind. If the visual cortex on one side is removed subjects become relatively blind in the contra-lateral hemifield. One of the most revealing studies of blindsight is Marcel's 1998 paper: " Blindsight and shape perception: deficit of visual consciousness or visual function?".

It is useful when considering blindsight to contemplate for a while the appearance of the world with both eyes closed and then with one eye closed. When both eyes are closed our experience is of a darkish space radiating out from our heads, with one eye closed we tend

to ignore the darkish areas that cannot be seen even though they are still present. Marcel notes that patients who have a right blind field still have an underlying visual field on the right side and that this can even contain conscious visual experience. This sounds a bit like the darkish space that we all experience if deprived of visual input on one side. As Marcel says: "A question that naturally arises is whether the loss is a 'total' loss of visual consciousness in the blind field. It is often assumed to be so, especially by those who discuss blindsight without carefully reading the literature or working with the subjects. One can immediately respond negatively to the question.."

The consciousness of the completion of Kanizsa figures in blindsight patients is particularly indicative of the preservation of the field even though the content was largely missing. A Kanizsa figure is shown below:

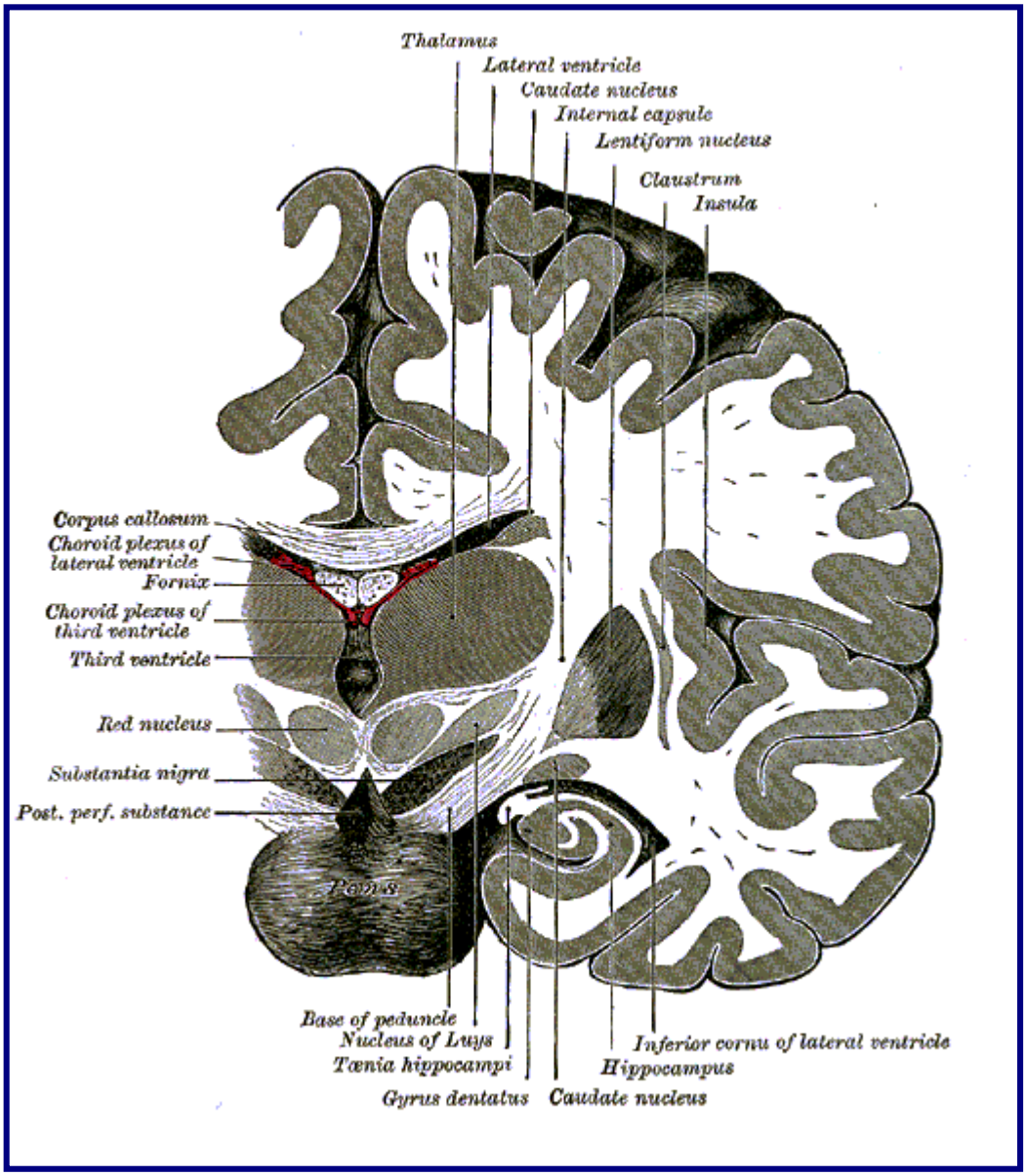


Blindsight patients can see the triangle completed in the blind field.

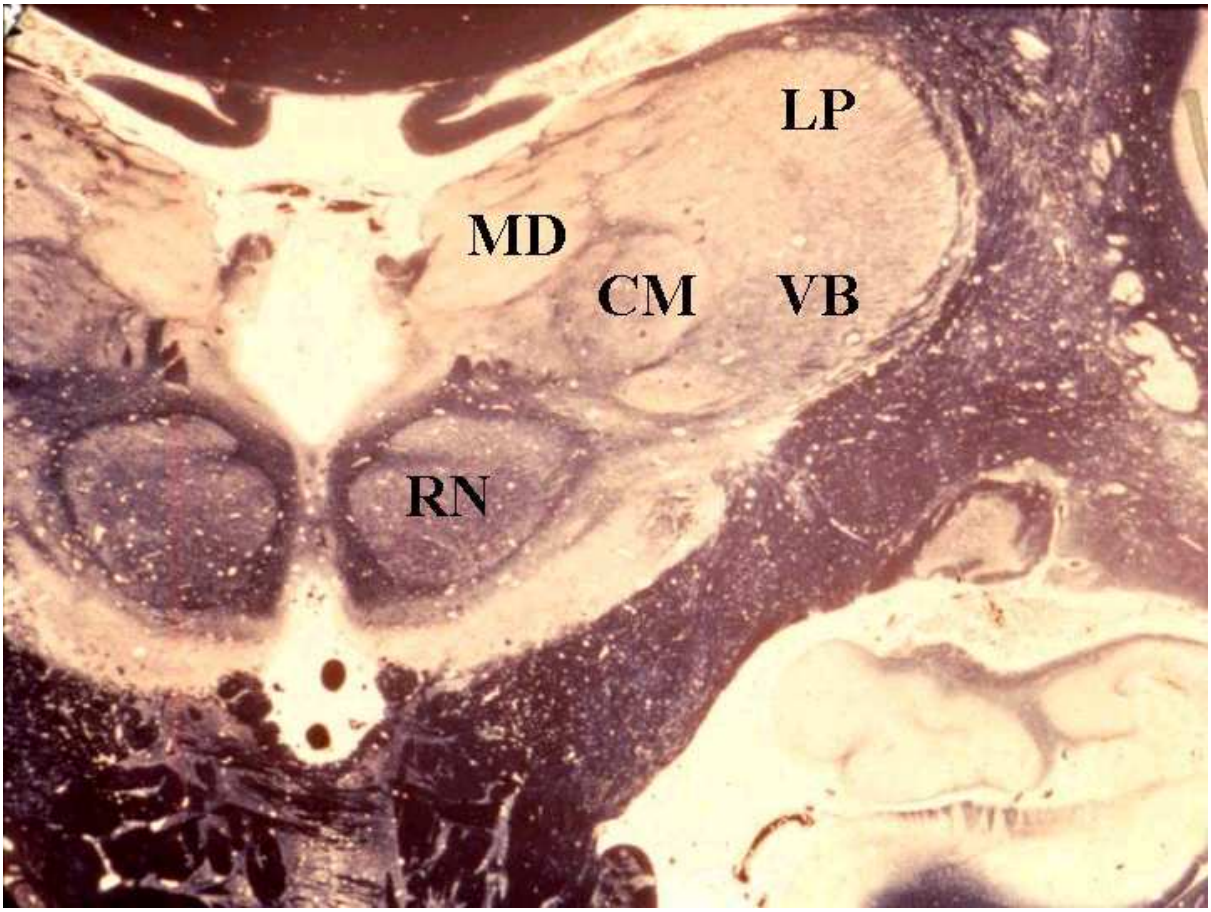
If we put Marcel's observations together with cortical anatomy and function it seems that the space of our experience is located outside of the cerebral cortex. The cortex generates much of visual and other content but it does not generate the space. A reasonable hypothesis is that the field of brain activity that is the space of our conscious experience is located in the sub-cortical brain. This space is loaded with the output of the cortex.

The Role of the Thalamus

The thalamus is connected to the entire bottom layer of the cerebral cortex. It is the nexus of the various cortical processors as well as a recipient of independent input from most of the rest of the brain.



The thalamus is subdivided into numerous small and medium sized nuclei that between them receive inputs from every process in the nervous system (the white fibres in the illustration above largely penetrate the thalamus). The thalamic nuclei are interconnected which means that any of them could, potentially host activity from anywhere in the body or brain. Although the founders of neurology such as Hughlings Jackson and Penfield & Jasper located conscious experience in the diencephalon, including the thalamus, this is no longer the conventional wisdom. The small size of the thalamic nuclei means that they cannot support the processes that are assumed to compose access consciousness, however, even some of the smallest thalamic nuclei host millions of synapses so size would not be an obstacle if the thalamus contains the substrate of phenomenal consciousness. Indeed, the diencephalon and the thalamus in particular can be shown to be excellent candidates for a possible location of phenomenal experience.



The Intralaminar Nuclei of the thalamus. The white space above and to the left of RN is the third ventricle. MD=mediodorsal nucleus. CM=Centromedian nucleus, RN=red nucleus (not part of thalamus) The black areas are stained white fibres. Picture from: http://www.neurophys.wisc.edu/ University of Wisconsin and Michigan State Comparative Mammalian Brain Collections. Preparation of image has been funded by the National Science Foundation, as well as by the National Institutes of Health. May only be used with these acknowledgements.

If the thalamus contains a location for conscious experience then lesions should abolish this experience. Unlike the cerebral hemispheres, lesions of the thalamus do indeed seem to abolish consciousness. The area that is most sensitive to lesions contains the Intralaminar Nuclei, especially the Parafascicular and Centromedian Nuclei. If these are damaged bilaterally patients suffer death, coma, akinetic mutism, hypersomnia, dementia and other

equally serious impairments of consciousness that depend upon the size and placement of the lesions (Bogen 1995, Schiff & Plum 1999). In cases of fatal familial insomnia, in which patients exhibit many of these symptoms, there is marked neuron loss in the Intralaminar Nuclei (Budka 1998). The symptoms of bilateral damage to the ILN are often so severe that it is possible that the patients cease to be conscious and are being coordinated by automatic cortical processes.

Laureys et al (2002) investigated recovery from 'persistent vegetative state' (wakefulness without awareness). They found that overall cortical metabolism remained almost constant during recovery but that the metabolism in the prefrontal and association cortices became correlated with thalamic ILN and precuneus activity. Again confirming that thalamo-cortico-thalamic activity is required for consciousness and that cortical activity by itself is not conscious. Yamamoto et al (2005) investigated persistent vegetative state and found that deep brain stimulation (25Hz) of the centromedian-parafascicular complex (19 cases) or mesencephalic reticular formation (2 cases) resulted in 8 of the patients emerging from persistent vegetative state.

As Bogen(1995) demonstrates, the ILN receive inputs, either directly or indirectly, from every part of the CNS but what do they do?

Interest in the thalamus has recently been revived by the theories of Newman & Baars (1993), Baars, Newman, & Taylor1998) and Crick & Koch (1990). In Baars, Newman and Taylors' (1998) theory it is suggested that "The brain stem-thalamocortical axis supports the state, but not the detailed contents of consciousness, which are produced by cortex". They also propose that the "nucleus reticularis thalami" (Thalamic Reticular Nucleus, TRN), which is a thin sheet of neurons that covers the thalamus, is involved in a selective attention system. This concept is reinforced by the way that point stimulation of the TRN causes focal activity in the overlying cortex (MacDonald et al 1998) and the way the TRN is organised topographically (ie: has activity that is like an electrical image of receptor fields).

The thalamus is ideally placed for integrating brain activity, if tiny parts of the thalamus are removed consciousness is abolished and the thalamus is involved in attention and the global integration of cortical activity. Any impartial judge might pronounce that the site of conscious experience has been found, possibly in the ILN of the thalamus, but no one can say how it works.

General Anaesthesia and the Thalamus

General anaesthesia should result in a profound depression of activity in the ILN if these are indeed the sites of the conscious state. White & Alkire (2003) administered halothane or isoflurane to volunteers and used positron emission tomography (PET) to monitor brain activity. They found severe depression of activity in the thalamus. The depression appeared to be higher in the non-specific nuclei than in the relay nuclei of the thalamus. In other words the anaesthesia is neither turning off the cortex nor turning off the input to the cortex but it is turning off an important part of the thalamus. Fiset et al (1999) have also demonstrated a similar pattern of medial thalamic inactivity and cortical activity in propofol anaesthesia. Suppression of cortical activity is not the cause of unconsciousness; for instance, the anaesthetic agent chloralose leads to increased neural activity in the cortex relative to conscious patients (Cariani 2000).

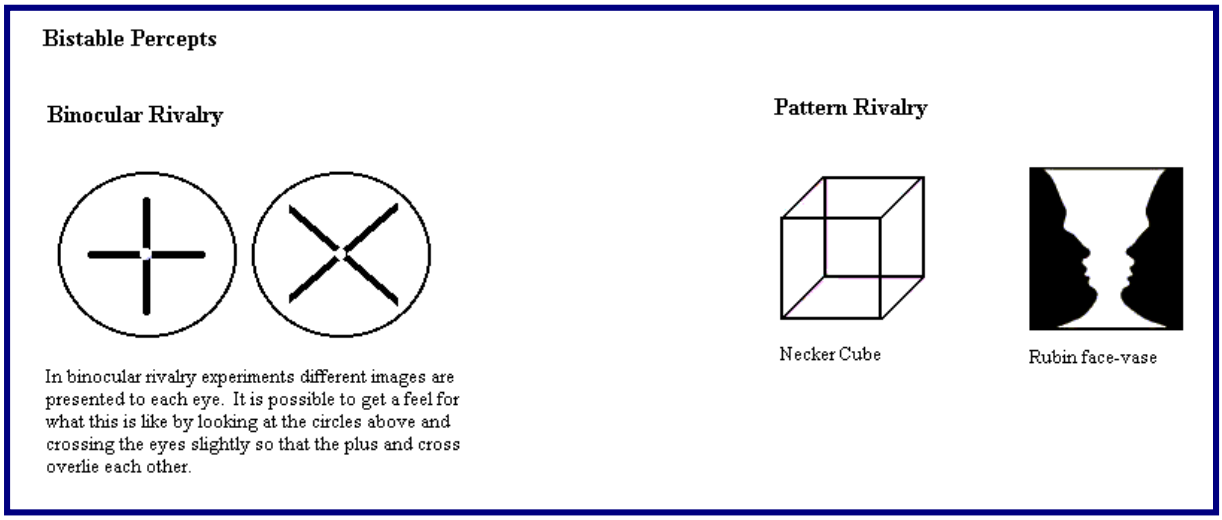
The function of consciousness

When we walk our conscious experience does not contain data about the control of the spinal, cerebellar and vestibular reflexes that keep us on an even keel. When we reach out for a cup our conscious experience only contains data related to the need for the cup, not data about the elaborate control system that enables the action. When we talk the words just come into mind, we do not painstakingly control the syntax and vocal chords. When our attention shifts the conscious experience containing the shift happens after the attention has shifted. This passive nature of experience recurs throughout the neuroscience of consciousness from the "readiness potential" to the "auditory continuity illusion". So what does conscious observation do? The medical evidence of the lack of consciousness in some forms of delirium, mutism, PVS etc. suggest that the role of conscious observation is to stabilise the brain so that it acts as a coordinated whole. Conscious observation is an orderly arrangement of events, a stable groundform that reflects the environment and composes the stage for action. It could be speculated that if quantum events were prominent in brain function then such a groundform would be essential but even a classical brain might require a stabilising form that could be continuously compared with the world beyond the body.

Rivalries, Synchronisation and Workspaces

Binocular Rivalry, Pattern Rivalry and Binocular Fusion

Sir Charles Wheatstone (1838) was the first scientist to systematically investigate binocular rivalry. Binocular rivalry occurs when different images are presented to the left and right eyes. The subject sees successively one image, a combined image and then the other image. The swapping of images can take a second or more. Binocular rivalry is of interest in consciousness research because the parts of the brain that contain the dominant image should also be those parts that are contributing to conscious experience. Binocular rivalry involves at least two components; the first switches from one image to a merged image and then to the other image and the second permits the view to be part of conscious experience.



The switching of one image for another may involve selecting one of the images as the percept or selecting one of the eyes. Blake et al (1979) performed an experiment in which subjects could change the image at a given eye by pressing a button. When a particular image became dominant they pressed a button to change the image at the eye receiving the dominant image for the non-dominant image. They found that the subjects immediately experienced the second image as the dominant image. This suggests that binocular rivalry is selecting between eyes rather than images. Lehky in 1988 proposed that the switching may be occurring as a result of feedback between visual cortical area V1 and the Lateral Geniculate Nucleus (a thalamic relay - see Carandini et al 2002) and Blake in 1989 also proposed that the switching occurred at the level of area V1. (Visual cortical area V1 receives visual input direct from the LGN.)

Tong (2001) has argued that, in humans, the switching of images in binocular rivalry may occur at the earliest levels in the visual cortex. In particular, Tong and Engel (2001) used an elegant technique measuring the activity in the visual cortex that represents the blind spot of the eye to show that almost complete switching to the dominant image occurs at the level of visual cortical area V1. In support of this idea of switching at the level of V1 or even before the cortex, Kreimann et al (2001, 2002) used direct electrode recordings in human cortex and found that the activity of most neurons changed with the percept. Other experiments

have not shown a single locus in the brain where the suppressed sensory information gets switched out (Blake & Logothetis 2002, Leopold & Logothetis 1996, Gail et al., 2004).

Functional MRI has also shown cortical activity outside of sensory visual cortex related to both images in binocular rivalry. Lumer et al (1998) found that only the fronto-parietal areas of cortex switched with the percept, Fang & He (2005) found that activity relating to both suppressed and unsuppressed images were present in the dorsal stream of the visual system. Wunderlich et al (2005) and Haynes et al (2005) have both found suppression at the level of the lateral geniculate nucleus using fMRI in humans.

Pasley et al (2004) have shown that, even during suppression, fearful faces can produce activity in the amygdala (see Pessoa (2005) for a review).

Rivalry alternations seem to be the result of widespread activity changes that cover large parts of the brain, including but not necessarily originating at the earliest sensory stages of visual processing. Most investigators have found that, once switching has occurred, there are areas of the brain that contain activity that is solely related to the percept but this varies from most of the cortex to largely more frontal regions depending upon the study. The most likely explanation for binocular rivalry is that the switching occurs at the level of the LGN as a result of feedback from the cortex.

Pattern Rivalry is also of interest in consciousness research for the same reasons as binocular rivalry. In pattern rivalry a figure may have two or more forms that replace each other. Typical examples of such figures are the Necker cube and Rubin's face-vase. The similarity of the time course of the switching between percepts in binocular rivalry and pattern rivalry has led many authors to suggest that these involve the same mechanism. Logothetis et al (1996) used novel dichoptic stimuli (different images to each eyes) to produce a form of rivalry that seems to involve switching at levels in the cerebral cortex that are more distal to the sensory stimulus than V1. Leopold and Logothetis (1999), on the basis of their work with monkeys, state that "..many neurons throughout the visual system, both monocular and binocular, continue to respond to a stimulus even when it is perceptually suppressed.". Kleinschmidt et al (1998) investigated pattern rivalry with MRI and found activity in higher order visual areas during change of dominant pattern. Pettigrew (2001) also describes effects on rivalry due to thought and mood that may require involvement of large areas of cortex in the switching operation and stresses the way that V1 represents different visual fields in different hemispheres of the brain so that inter-hemispheric switching must also be considered.

It seems likely that the change of dominant pattern or percept is associated with higher level cortical activity but once the dominant percept is established many of the visually responsive neurons in the cortex are switched over to the new percept. This might account for the similarities in timing of binocular and pattern rivalry and the disparate results found by the various groups of authors. In the words of Kleinschmidt et al (1998):

"The transient activity fluctuations we found suggest that perceptual metastability elicited by ambiguous stimuli is associated with rapid redistributions of neural activity between separate specialized cortical and subcortical structures."

Which permits both the idea of selecting particular eyes or percepts, perhaps by feedback that switches a thalamic relay on the basis of cortical processing of patterns. Once the cortex

has switched the thalamic relay most of the neurons in V1 would become exposed to the dominant percept but there would still be a few neurons in the cortical visual system receiving data from the non dominant image.

The investigations of binocular and pattern rivalry provide evidence that conscious visual experience is probably distal to V1 (ie: cortex or thalamus).

Perceptual rivalry may be part of complex decision making rather than being simply a switch to blank out unwelcome input. It is clear from the Rubin face-vase that pattern rivalry is linked to recognition and would involve a complex delineation of forms within cortical processing. This would suggest that many areas of cortex should be involved before a particular percept is made dominant. Pettigrew (2001) argues that rivalry is the result of a complex phenomenon rather than being simply a switching event. Pettigrew's discovery that laughter abolishes rivalry also points to a complex cortical system for switching percepts. Pettigrew proposes that complex cortical processes control rivalry and that the actual switching of percepts is performed sub-cortically in the Ventral Tegmental Area. He concludes his review of the problem by noting that "Rivalry may thus reflect fundamental aspects of perceptual decision making.." Pettigrew (2001).

Another effect, known as "binocular fusion", provides further compelling evidence for the non-conscious nature of the cerebral cortex. In binocular fusion images from both eyes are fused together to create a single image in experience. Moutoussis and Zeki (2002) used a form of binocular fusion in which images of faces were flashed at 100ms intervals to both eyes simultaneously. When both eyes received images of the same colour the subject could see the faces but when one eye received a green image on a red background and the other a red image on a green background the subjects reported seeing a uniform yellow field that contained no faces.

fMRI scans of the subject's brains showed that when both eyes were exposed to images of the same colour the part of the brain that deals with faces was active and when each eye received images of different colours the same areas of brain showed activity. In other words the cortex contained strong activity related to faces whether or not faces were experienced. Moutoussis and Zeki found a similar effect when they used images of houses instead of images of faces. The authors concluded that: "The present study further suggests that there are no separate processing and perceptual areas but rather that the same cortical regions are involved in both the processing and, when certain levels of activation are reached and probably in combination with the activation of other areas as well, the generation of a conscious visual percept".

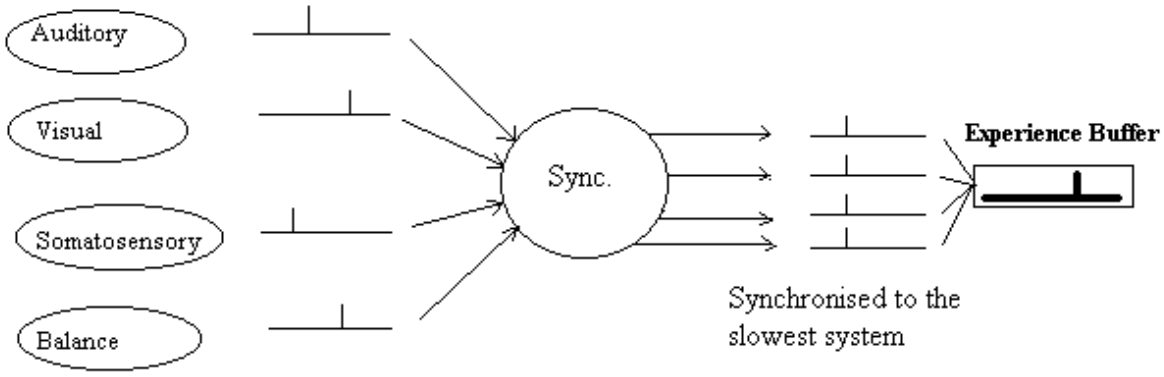
This conclusion does not seem to be supported by the data. There is no evidence that any area of cortex contains the percept itself. The experiment shows that the cortex contains data relating to both red and green faces which suggests that the cortex is not the site of the conscious percept. The percept is most likely distal to the cortex in the thalamus.

It is interesting that Fries et al (1997) found that neurons that were activated by the dominant image in binocular rivalry fired synchronously whereas those that were activated by the non-dominant image did not. Thalamocorticothalamic oscillations are the most likely source for synchronising neurons over whole areas of cortex, again suggesting that the conscious percept is located in the thalamus rather than the cortex.

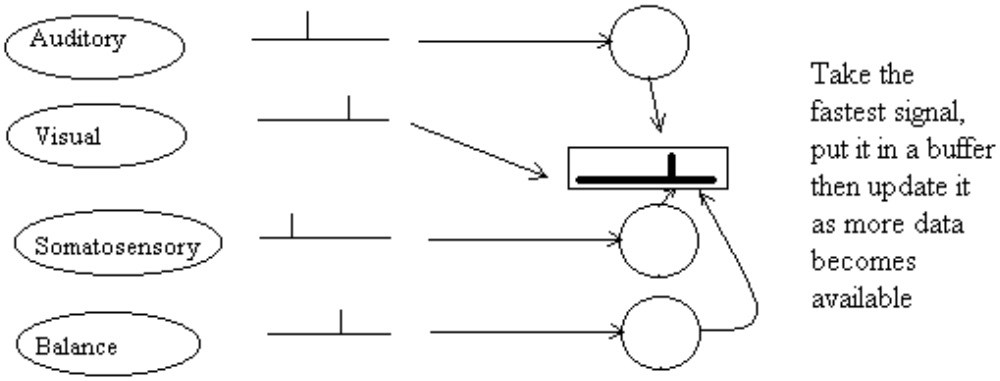
Synchronisation of Neural Processes

Our experience seems to contain entities with their attributes attached to them at the correct places in space and time. When a dog barks we see its jaws open at the same time as the bark and both jaws and bark are at the same location. We take this for granted but the brain must be engaging in some complex processing to achieve this synchronised and appropriately positioned set of objects and events. The illustration below shows the two basic processes that might be used to synchronise events between the different specialised processors in the cerebral cortex and brain in general.

Synchronising Multiple Processes - Experiencing a Stimulus



Option 1: Wait till all data has arrived then integrate it.



Option 2: Take the fastest then update it

These options are not exclusive: a mixture of 1 & 2 might occur.

In the first option a complete model of sensation, dream etc. may be created and then allowed to become part of conscious experience. In the second model events are released into experience as fast as possible but are synchronous when recalled, having been synchronised in a storage buffer. There is a third option in which there is no synchronisation of events so that the output from different processors would occur at different times.

The 'experience buffer' would be a volume of brain in which a succession of events could be recorded. The buffer might either be updated in steps, the previous content being discarded, or continuously updated with the oldest content being lost continuously.

In the first option events from different processes would always appear to be simultaneous unless the experience buffer were updated as a series of steps in which case any changes at around the moment of updating might appear in successive buffers. For instance, if change of position were processed before change in colour a circle on a screen that changed from green to red at the start of a motion might seem to be briefly green during the motion and then turn red.

In the second model events from different processors might appear asynchronous at the moment of experience but synchronous when recalled.

Colour vision and motion vision are processed in different parts of the visual cortex and in distinct parts of visual cortical areas V1 and V2. They are different processes and hence ideal for studying the synchronisation of cortical activity. Moutoussis and Zeki (1997) presented subjects with moving coloured squares on a computer screen that changed from red to green or vice versa as they changed direction of movement. It was found that subjects seemed to perceive changes in colour some 70-80 msecs before they perceived a change in the direction of motion of the squares. Further work by Arnold et al (2001) and Arnold and Clifford (2001) have confirmed that colour changes seem to be perceived before motion. Arnold and Clifford (2001) also found a quantitative relationship between the colour/motion asynchrony and the direction of change of motion, complete reversals of direction giving rise to the greatest asynchrony between the detection of colour and motion changes.

Moutoussis and Zeki (1997) conclude by stating that the asynchrony of neural processes shows that "..the perception of each attribute is solely the result of the activity in the specialised system involved in its processing..". It seems more likely that the experiments simply show that slow neural processes are not synchronised before they become percepts (the third option above). The experiments are excellent evidence for the concept of the cortex as a set of specialised processors that deliver their output asynchronously to some other place where the output becomes a percept.

These experiments on colour and motion suggest that there is no synchronisation between the processes that deal with these two aspects of vision. Another set of experiments by Clifford et al (2003) supports this idea of processing being asynchronous. They asked subjects to perform a variety of judgements of when visual events occurred and found that the degree of synchrony of one visual event with another depends on the type of judgement. Different judgements probably use processors in different areas of cortex and the output from these arrives asynchronously at the part of the brain that supports the percept.

When the percept is formed there must be feedback to the cortical processes that create its content. Otherwise it would not be possible to report about the percept and the cortex would be unable to direct processing to the percept in preference to other, non-conscious cortical data.

Although slow processes (20 milliseconds to 1 second) do not seem to be synchronised there is some evidence for very rapid synchronisation. Andrews et al (1996) revisited a problem raised by the famous physiologist Charles Sherrington. Sherrington considered the

phenomenon of 'flicker fusion' in which a flickering light appears to be a continuous steady light if it flashes on and off at frequencies of about 45 Hz or higher. He reasoned that if the images from both eyes are brought together to form a single image then the frequency at which a flickering light appears to be steady should depend on whether one or two eyes are used. Flicker fusion should occur if each eye receives alternate flashes at only half the normal flicker fusion frequency. The flicker should disappear if the left eye receives flashes at 23 pulses per second and the right eye receives alternate flashes at 23 pulses per second. When Sherrington performed the experiment he found that this was not the case, using approximate figures, each eye required 46 pulses per second for fusion to occur. Sherrington proposed that the flicker fusion in alternate binocular presentation was occurring "psychically", outside of normal physiological processes.

Andrews et al duplicated Sherrington's result but investigated it further. They found that when lights were flashed in each eye alternately at low frequences (2 Hz) the experience was the same as a light being flashed in both eyes at this rate. At frequencies of four Hz and higher the subjects began to report that the lights being flashed alternately in both eyes seemed to flicker at the same rate as lights being flashed in both eyes at half the frequency. It seemed as if a flash in one eye followed by a flash in the other eye was being perceived as a single flash or "conflated" as the authors put it. The authors explained this effect by suggesting that the brain activity corresponding to the flashes was sampled for a short period and any number of flashes occurring during this period became perceived as a single flash. The maximum rate of sampling would be about 45 Hz. This idea is similar to option (1) above, where the buffer is filled and emptied 40 - 50 times a second.

An experience buffer that is refreshed at 40-50 times a second might also explain the results obtained with colour and motion asynchrony because synchronisation between processes may well happen too quickly to affect processes that occur at very slow rates. Singer and Gray (1995), Singer (2001) have proposed that synchronisation between neurones at about 45Hz is the discriminator between those neurones with activity that contributes to conscious experience and activity in other neurones. A rapid refresh rate in a sychronising buffer agrees with the results found by Fries et al (1997) in which visual cortical neurones that represent a percept underwent synchronous oscillations in the gamma frequency range (39-63 Hz). Tononi et al (1998) have also found synchronisation of neural activity in neurones that represent the percept.

The gamma frequency oscillations are intrinsic to the cortex but are triggered by the thalamus and are part of the 'arousal system'. Readers should be wary of the term 'arousal system' because it evokes the idea of something waking up a conscious cortex. The cortex can be fully active during sleep and even during pathological unconsciousness such as persistent vegetative state so it is possible that the arousal centres themselves or nearby structures actually host phenomenal consciousness.

EEG and synchronisation

If electrodes are placed on the scalp varying electrical potentials of a few tens of microvolts can be recorded between the electrodes. Recordings of potentials from electrodes on the scalp are known as electroencephalograms (EEGs).

The potentials recorded in the EEG are due to postsynaptic potentials in nerve cells. The

EEG is insensitive to the activity of single cells and occurs as a result of relatively slow, **synchronised**, changes in large areas of cells. The differences in potential between two scalp electrodes are largely due to depolarisation and hyperpolarisation of the dendritic trees of cortical pyramidal cells. The folding of the cortex (gyri) is problematical for recording and interpreting EEGs because opposing layers of cortex can cancel any net potentials.

The EEG shows rhythmic activity. This is conventionally divided into the following frequency bands:

Delta waves 0-4 Hz

Theta waves 4-8 Hz

Alpha waves 8-12 Hz

Beta waves >10 Hz

Gamma waves (also called fast beta) 25-100 Hz

EEGs also contain short bursts of activity called spindles and very fast oscillations (VFOs). Spindles last for 1-2 seconds and contain rhythmic activity at 7-14 Hz. They are associated with the onset of sleep. The VFOs consist of short bursts at frequencies of over 80 Hz.

When the eyes are closed the amplitude of activity from most pairs of electrodes is increased compared with when the eyes are open. When subjects are awake the EEG consists mainly of alpha and beta activity with considerable low amplitude gamma when the eyes are open. In stage 1 sleep the EEG consists of theta waves, in stage 2 sleep of varied activity and spindles, in stage 4 sleep of delta and during REM sleep of beta and theta activity. In epileptic seizures there tends to be high amplitude activity with pronounced synchronisation between many pairs of electrodes.

The rhythmic electrical activity is due to cortical feeback loops, cortico-cortical synchronisation, thalamic pacemakers and thalamo-cortical synchronisation. VFOs have been attributed to the activity of electrical connections between cells (dendro-dendritic gap junctions) (Traub (2003)).

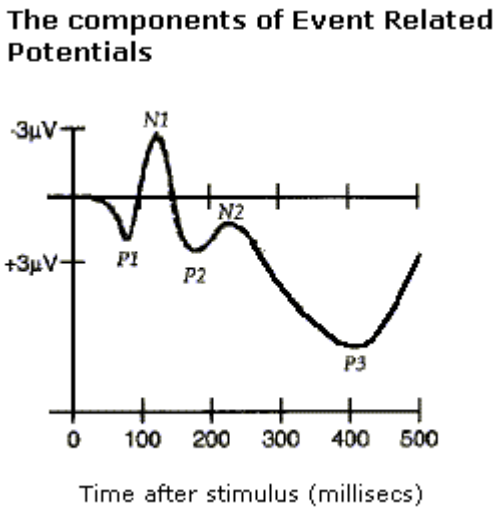
The gamma activity, centred on a frequency of 40 Hz appears to be related to activity in cortical interneurons that form electrical connections between their dendrites (Tamas et al 2000). These oscillations can be triggered by high frequency stimulation of the posterior intralaminar nuclei of the thalamus (Barth and MacDonald 1996, Sukov and Barth 2001) and as a result of activation of the reticular system (Munk et al 1996). This suggests that stimulation of cortex by thalamic sensory relays triggers gamma band activity in the cortex. A shift from gamma to beta waves can occur in human event related potentials after about 0.2 secs (Pantev 1995, Traub et al 1999).

The alpha activity is related to thalamic pacemakers, perhaps as a result of intrinsic oscillatory activity in thalamic sensory relays (see Roy & Prichep 2005 for a brief review). Theta activity, which occurs during some cognitive tasks and mental arithmetic involves a loop from the cortex to the non-specific thalamic nuclei. Delta activity seems to be endogenous to cortex when input is suppressed during sleep. Beta activity is due to cortico-cortical interactions, often after a brief period of gamma activation. It should be noted that

gamma and beta activity can be expressed as impulses in cortico-thalamic pathways and that when cortical and thalamic activity is correlated there is a conscious state. In other words gamma or beta waves in the cortex are not correlates of consciousness on their own - see for instance Laureys et al (2002).

Event related potentials

After a sudden event there are a characteristic set of changes in EEG activity known as **event related potentials** or ERPs. The time course of the ERP is shown in the diagram below.



ERPs occur in response to novel stimuli and are also produced by brief transcranial magnetic stimulation (TMS)(Iramina et al 2002). The slow component is known as the P3 or P300 phase of the ERP. It is due to activation of areas of the brain that are relatively remote from the primary sensory areas of brain.

Nieuwenhuis et al (2005) have reviewed the origin of the P300 ERP: "To summarize, convergent evidence suggests that P3-like activity can be recorded in several, widely separated brain areas. These include some medial temporal and subcortical structures (e.g., the hippocampal formation, amygdala, and thalamus), but these structures are unlikely to contribute directly to the scalp-recorded P3.". According to Nieuwenhuis et al (2005) the recorded P300 may be due to temporo-parietal and prefrontal cortical activity. Linden (2005) has also concluded that widespread, but specific, cortical activation is correlated with the recorded P300 ERP.

The generator of the P300 is still obscure. Nieuwenhuis et al (2005) consider that the Locus coeruleus, a nucleus in the pons that regulates task related attention and part of the sleep-wake cycle, may be responsible. In line with this, Mashour et al (2005) have discovered that TMS induced P300 activity is reduced in unconscious states.

Whether the P300 is related to Libet's 0.5 second delay is still obscure but the discovery that the P300 occurs in association with subliminal stimuli (stimuli that do not enter

awareness)(Bernat et al 2001) suggests that it is associated with non-conscious cortical processing. Williams et al (2004), in an investigation of subliminal and supraliminal fear perception, found that "conscious fear perception was distinguished by a more prominent N4, peaking around 400 msec"; the N4 component follows the P300 component in the succession of phases of the ERP. Williams et al considered that the earlier phases in the ERP are probably related to non-conscious processing. In contrast Vogel et al (1998) found that suppression of the P300 was associated with suppression of awareness.

Global Workspace Theory

Global Workspace Theory is the idea that somewhere in the brain there is a facility that integrates the processes that occur in the various separate areas of the brain. The theory was first proposed by Descartes as the *sensus communis*, the common sense, but the modern form of the theory dispenses with the idea of a point soul looking at the brain. In modern Global Workspace theory it is proposed that an area of brain receives input from most of the cerebral cortex and provides integrative output that can solve problems across sensory and pre-motor modalities.

Modern Global Workspace Theory has been championed by Baars (1983, 1988).

There is increasing evidence for a Global Workspace or Global Workspaces in the brain. Much of this evidence comes from fMRI, single unit and magnetoencephalography studies in which it is shown that non-conscious or subliminal processing mainly occupies primary, sensory cortex whereas conscious processing occupies large areas of cerbral cortex.

In binocular rivalry the stimulus that is consciously perceived is responsible for relatively intense activation of large areas of brain whereas the non-conscious stimulus is often suppressed (see above and Sheinberg & Logothetis (1997), Tononi et al (1998)). The suppression is likely to occur in the Lateral Geniculate Nuclei which suggests a role for the Thalamic Reticular Nuclei, which modulate LGN activity, in the control of the percept.

Masking and visual awareness

Word masking has also been used to investigate the idea of a Global Workspace. When a word is presented on its own for a few tens of milliseconds it remains readable but if it is immediately succeeded by, or accompanied by, another word it becomes indistinct or invisible. This effect is known as "word masking". Vogel et al (1998) have investigated a version of word masking known as the "attentional blink". They found that when stimuli became invisible the P3 component of the Event Related Potential, which peaks at around 300-500 millisecs after a stimulus, was completely suppressed. The P3 component of the ERP has been related to the lodging of data in working memory and also to gamma band activity in the EEG. This strongly suggests the involvement of a cortico-thalamic loop in the "attentional blink". The delay of 0.3 to 0.5 secs is typical of the time required for conscious awareness (see above).

Word masking in conjunction with fMRI and Event Related Potential (ERP) recordings has been used by Dehaene et al (2001) to expose control by a central mechanism. It was found that masked words activate mainly the visual cortex and ventral stream (inferior temporal lobe) whereas visible words also activated distant parietal, prefrontal and cingulate sites.

Dehaene et al (2003) and found that the dynamics of the loss of visibility of words in an attentional blink experiment could be modelled by a simulated cortico-thalamic loop. In their simulation a distributed cortical process determined which events would receive attention and the system used the thalamic gating systems to exclude those that did not receive attention.

Tse et al (2005) have used purely visual stimuli in masking experiments and concluded that, in the case of purely visual stimuli, the neural correlates of awareness were limited to the occipital cortex:

"We suggest that there are both lower and upper bounds within the visual hierarchy for the processing of visual masking and the maintenance of visual awareness of simple unattended targets; the lower bound is at least as high as the border between V2 and V3, and the upper bound is within the occipital lobe, possibly somewhere downstream of V4."

This discovery would mean that activation of large areas of cortex are unnecessary for awareness.

Attention and the global workspace

Baars (2002) in his review of evidence for the Global Workspace Theory quotes many other experiments that show activation of larger areas of cortex in response to conscious stimuli compared with unconscious or subliminal stimuli.The effect is also seen in change blindness, learning and attention. Newman and Baars (1993) consider that the "workspace" is fairly global in the brain:

"This Neural Global Workspace (NGW) model views conscious processes in terms of a globally integrative brain system. The neural circuitry contributing to this system is not only widely distributed across the neocortex, but includes key corticothalamic and midbrain circuits as well. These cortico-subcortical circuits are hypothesized to be critical to understanding the mechanisms of attentional control that provide an essential basis for the conscious processing of information".

However they focus particularly on the role of the thalamic Reticular Nucleus and cortico-thalamic connectivity in the control of attention.

Other ideas for the location of the Global Workspace are the idea of Singer et al. that gamma synchrony controls access to the content of consciousness and Llinas et al. (1998) that the thalamus is the hub through which communication occurs between areas of cortex.

One of the problems with Global Workspace theory is that it suggests that attention, working memory, cognitive control and consciousness may all be in the same area of the brain. It is likely that the mechanisms of attention, working memory, and cognitive control may involve several, interlinked systems perhaps co-opting the basal ganglia in the process. In view of this Maia and Cleeremans (2005) propose that ".. attention, working memory, cognitive control and consciousness are not distinct functions implemented by separate brain systems. Attempting to find separate neural correlates for each may therefore be the wrong approach. Instead, we suggest that they should be understood in terms of the dynamics of global competition, with biasing from PFC (prefrontal cortex).". The inclusion by Maia and Cleeremans of consciousness with distributed attention, working memory and cognitive control is reminiscent of Zeki & Bartel's idea of microconsciousness.

It should be noted that, in common with Libet's data, the percept seems to be available to phenomenal consciousness some 0.3 to 0.5 secs after a stimulus; this suggests that whatever determines the content of phenomenal consciousness operates before events become part of phenomenal consciousness. This relegates phenomenal consciousness from being a controller of attention to being the recipient of content that is the subject of attention. This finding is consistent with the philosophical problem of the **apparently** epiphenomenal nature of phenomenal consciousness.

Given the data on the timing of conscious awareness it seems that there may be two "workspaces", an active workspace that models the world, discarding and suppressing data during rivalry, and a passive workspace that receives the final, edited product. The active workspace would correlate with the cortical systems stressed by Dehaene et al and Maia and Cleermans although, given the results of Tse et al., the workspace would be limited to small zones of cortex. The loading of the passive workspace with the output of the active workspace would correlate with thalamo-cortical activity during component P3 of the ERP in which data is transferred from the cortex to the thalamus. This workspace might constitute the source for reports of the content of phenomenal consciousness.

Llinas et al (1998) have proposed two parallel cortico-thalamic attentional systems, one of which is related to the thalamic specific nuclei and the other to the thalamic non-specific nuclei, especially the ILN. The non-specific system would be related to consciousness itself.

Bibliography

Defining the States of Consciousness. Tassi, P., Muzet, A. (2001) Neuroscience and Behavioural Reviews 25(2001) 175-191.

Neuroanatomy: "Digital Slice of Life" by Stensaas and Millhouse

http://medlib.med.utah.edu/kw/sol/sss/subj2.html

See also: http://medlib.med.utah.edu/kw/sol/sss/index.html

EEG's: Coles, M.G.H., Rugg, M.D. Event Related Brain Potentials: An Introduction.

http://whalen.psych.udel.edu/667/1.What_is_ERP/ColesRugg1995chpt1.pdf

Visual System: Tong, F. (2002). Primary Visual Cortex and Visual Awareness. Nature Reviews Neuroscience 4, 219 (2003)

http://www.psy.vanderbilt.edu/tonglab/publications/Tong_NRN2003.pdf

Professor Bogen's Consciousness Page

http://www.its.caltech.edu/~jbogen/text/toccons.htm

Demonstrations of Auditory Illusions and Tricks. Yoshitaka Nakajima

http://www.kyushu-id.ac.jp/~ynhome/ENG/Demo/illusions2nd.html

References

- Aleman, A., Ramsey, N., Van Honk, J.E., Kessels, R.P.C., Hoogduin, J., Schutter, D.L.G., Postma, A., Kahn, R.S., De Haan, E.H.F. (submitted). Parietal cortex mediates visuospatial imagery: evidence from fMRI and rTMS. http://igitur-archive.library.uu.nl/dissertations/1954455/c10.pdf

- Alkire, M.T., Haier, R.J., Fallon, J.H., and Barker, S.J.(1996). PET imaging of conscious and unconscious verbal memory. Journal of Consciousness Studies, 3 (5-6) 1996 , pp. 448-462 http://www.imprint.co.uk/jcs_3_5-6.html

- Anderson, J. & Anderson, B. (1993). "The Myth of Persistence of Vision Revisited," Journal of Film and Video, Vol. 45, No. 1 (Spring 1993): 3-12. http://www.uca.edu/org/ccsmi/ccsmi/classicwork/Myth%20Revisited.htm

- Andrews, T.J., White, L.E., Binder, D. & Purves, D. (1996). Temporal events in cyclopean vision. Proc. Natl. Acad. Sci. USA Vol. 93, pp. 3689-3692, April 1996. Neurobiology. http://www.dur.ac.uk/tim.andrews/PNAS_93_3689.pdf

- Arnold, D.H., Clifford, C.W.G. (2001) Determinants of asynchronous processing in vision. Proceedings of the Royal Society of London, Series B: Biological Sciences, 269, 579 - 583. http://www.psych.usyd.edu.au/staff/colinc/Reprints/AC02.pdf

- Arnold, D.H., Clifford, C.W.G, Wenderoth, P. (2001). Asynchronous processing in vision: Color leads motion. Current Biology, 11, 596 - 600. http://www.psych.usyd.edu.au/staff/colinc/Reprints/ACW01.pdf

- Baars, B.J. (1983) Conscious contents provide the nervous system with coherent, global information. In Consciousness and Self-Regulation (Vol 3). Davidson, R.J. et al., editors. Plenum Press.

- Baars, B.J. (1988) A Cognitive Theory of Consciousness, Cambridge University Press.

- Baars, B., Newman, J., & Taylor, J. (1998) Neuronal mechanisms of consciousness: A Relational Global Workspace framework., in Hameroff, Stuart and Kaszniak, Al and Laukes, James, Eds. Toward a Science of Consciousness II: The second Tucson discussions and debates., chapter 24, pages 269-278. MIT Press. http://cogprints.ecs.soton.ac.uk/archive/00000946/

- Baars, B.J. (2002) The conscious access hypothesis: origins and recent evidence. Trends in Cognitive Sciences. Vol 6, No 1, 47-52. 2002. http://www.cs.memphis.edu/~wrcm/BaarsTICS2002.pdf

- Barth, D. S. and MacDonald, K. D. (1996) Thalamic modulation of high-frequency oscillating potentials in auditory cortex. Nature. 383, 78-81.

- Bernat E, Shevrin H, Snodgrass M. (2001). Subliminal visual oddball stimuli evoke a P300 component. Clin Neurophysiol. 2001 Jan;112(1):159-71.

- Blake, R., Westendorf, D. and Overton, R. (1979) What is suppressed during binocular rivalry? Perception 9: 223-231.

- Blake, R. and Lee, S.-H. (1999). Rival ideas about binocular rivalry. Vision Research. 39: 1447-1454 http://www.psy.vanderbilt.edu/faculty/blake/PDFs/Lee_Blake_99VR.pdf

- Bogen, J.E. (1995). On the neurophysiology of consciousness: I. An overview. Consciousness and Cognition, 4, 52-62. http://www.its.caltech.edu/~jbogen/text/bogen-95.pdf

- Bogen, J.E. (1995). On the neurophysiology of consciousness: Part II. Constraining the semantic problem. Consciousness and Cognition, 4, 137-158. http://www.its.caltech.edu/~jbogen/text/concog95.htm

- Bregman, A.S. (1990). Auditory scene analysis: The perceptual organization of sound. Cambridge, MA: MIT Press.

- Budka, H. (1998). Fatal Familial Insomnia Around the World. Brain Pathology, 8: 553-570. http://brainpath.medsch.ucla.edu/pdfs/0803pdf/budka.pdf

- Carandini M, Heeger DJ, & Senn W, (2002). A synaptic explanation of suppression in visual cortex, Journal of Neuroscience, 22:10053-10065, 2002. http://www.cns.nyu.edu/~david/ftp/reprints/Carandini-JNeurosci2002.pdf

- Cariani, P. (2000). Anesthesia, Neural Information Processing, And Conscious Awareness. Consciousness And Cognition Vol. 9, No. 3, September 2000 p. 387-395

- Castro-Alamancos, M.A., Connors, B.W. (1997). Thalamo-cortical synapses. Progress in Neurobiology, Vol.51, pp. 581 to 606, 1997 http://neurobio.drexel.edu/castroweb/progneurobio1.pdf

- Christmann, C. & Leuthold, H. (2004). The Quarterly Journal of Experimental Psychology: Section A. Volume 57, Number 2 / February 2004. 357-381.

- Clifford, C.W.G., Arnold, D.H., Pearson, J. (2003). A paradox of temporal perception revealed by a stimulus oscillating in colour and orientation. Vision Research 43 (2003) 2245-2253. http://www.psych.usyd.edu.au/staff/colinc/Reprints/CAP03.pdf

- Crick, F and Koch, C.(1990). Towards a neurobiological theory of consciousness Seminars in the Neurosciences (1990) 2, 263-275

- Crick, F and Koch, C.(1998). Consciousness and Neuroscience. Cerebral Cortex, 8:97-107, 1998 http://www.klab.caltech.edu/~koch/crick-koch-cc-97.html

- Culham, J.C., Dukelow, S.P., Vilis, T., Hassard, F.A., Gati, J.S., Menon, R.S. and Goodale, M.A. (1999). Recovery of fMRI Activation in Motion Area MT Following Storage of the Motion Aftereffect. The Journal of Neurophysiology Vol. 81 No. 1 January 1999, pp. 388-393. http://jn.physiology.org/cgi/content/full/81/1/388

- Deecke, L. (1996) Planning, preparation, execution, and imagery of volitional action, (Introduction / Editorial) in: L. Deecke W. Lang, A. Berthoz (Eds) Mental Representations of Motor Acts. Cognitive Brain Res 3 / Special Issue (2) 59-64. http://cogprints.ecs.soton.ac.uk/archive/00000015/00/COGBRAIN.html

- Dehaene, S., Sergent, C., and Changeux, J. (2003) A neuronal network model linking subjective reports and objective physiological data during conscious perception. PNAS July 8, 2003 vol. 100 no. 14 http://www.unicog.org/publications/DehaeneSergentChangeux_BlinkModel_PNAS2003.pdf

- Dehaene, S. et al. (2001) Cerebral mechanisms of word masking and unconscious repetition priming. Nat. Neurosci, 4, 752-758. http://www.unicog.org/publications/Dehaene_MaskedWords_NatureNeuroscience2001.pdf

- Destexhe, A. (2000). Modelling corticothalamic feedback and the gating of the thalamus by the cerebral cortex. J. Physiol (Paris), 94 (2000) 391-410. http://cns.iaf.cnrs-gif.fr/files/Brem2000.pdf

- Fang, F. & He, S. (2005).Cortical responses to invisible objects in the human dorsal and ventral pathways. Nat Neurosci. 2005 Oct;8(10):1380-5. Epub 2005 Sep 4.

- Ffytche, D. H., Howard, R. J., Brammer, M. J., David, A., Woodruff, P., & Williams, S. (1998). The anatomy of conscious vision: an fMRI study of visual hallucinations. Nature Neuroscience, 1, 738â€“742.

- Fiset, P., Paus, T., Daloze, T., Plourde, G., Meuret, P., Bonhomme, V., Hajj-Ali, N., Backman, S.B., and Evans, A.C. (1999). Brain Mechanisms of Propofol-Induced Loss of Consciousness in Humans: a Positron Emission Tomographic Study. The Journal of Neuroscience, July 1, 1999, 19(13):5506-5513. http://www.jneurosci.org/cgi/content/full/19/13/5506

- Fu, C.H.Y., Morgan, K., Suckling, J., Williams, S.C.R., Andrew, C., Vythelingum, G.N., McGuire, P.K. (2002). A Functional Magnetic Resonance Imaging Study of Overt Letter Verbal Fluency Using a Clustered Acquisition Sequence: Greater Anterior Cingulate Activation with Increased Task Demand. NeuroImage 17, 871- 879 (2002) http://www-bmu.psychiatry.cam.ac.uk/PUBLICATION_STORE/fu02fun.pdf

- Ganis, G., Thomson, W.L. and Kosslyn, S.M. (2004). Brain areas underlying visual mental imagery and visual perception: an fMRI study. Cognitive Brain Research 20 (2004) 226â€“241 http://www.wjh.harvard.edu/~kwn/Kosslyn_pdfs/2004Ganis_CogBrainRes20_BrainAreas.pdf

- Hultin, L. (1997). On Neuromagnetic Assessment of the Contingent Negative Variation. http://www2.lib.chalmers.se/cth/diss/doc/9697/HultinLeif.html

- Fries, P.. Roelfsema, P.R., Engel, A.K., KÃ¶nig, P. and Singer, W. (1997). Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. Proc. Natl. Acad. Sci. USA, Vol. 94, pp. 12699-12704, November 1997. http://www.pnas.org/cgi/content/full/94/23/12699

- Fuentealba, P., Crochet, S., Timofeev, I., Bazhenov, M., Sejnowski, T.J. and Steriade, M. (2004). European Journal of Neuroscience 20(1):111-119. http://www.theses.ulaval.ca/2004/22126/ch03.html

- Goaillard, J. & Vincent, P. (2002). Serotonin suppresses the slow afterhyperpolarization in rat intralaminar and midline thalamic neurones by activating 5-HT7 receptors. Journal of Physiology (2002), 541.2, pp. 453-465 http://jp.physoc.org/cgi/content/full/541/2/453

- Haynes, J., Deichmann, R. and Rees, G. (2005). Eye-specific effects of binocular rivalry in the human lateral geniculate nucleus. Nature 438, 496-499 (24 November 2005)

- Hughes S.W., Blethyn K.L. , Cope D.W. , & Crunelli V. (2003). Properties and origin of spikelets in thalamocortical neurones in vitro. Neuroscience. 2002;110(3):395-401

- Iramina, K., Maeno, T., Kowatari, Y. and Ueno, S.. (2002). Effects of Transcranial Magnetic Stimulation on EEG Activity .IEEE TRANSACTIONS ON MAGNETICS, VOL. 38, NO. 5, SEPTEMBER 2002

- Ishai, A., Ungerleider, L.G., Haxby, J.V.. (2000) Distributed neural systems for the generation of visual images. Neuron 28, 979-990.

- Johnson-Laird, P. N. (1980). Mental models in cognitive science. Cognitive Science, 4, 71-115. http://cognitrn.psych.indiana.edu/rgoldsto/cogsci/classics.html

- Keller, J. & Heckhausen, H. (1990). Readiness potentials preceding spontaneous motor acts: voluntary vs. involuntary control. Electroencephalography and Clinical Neuropsycholgy, 76, 351-361.

- King, C.C. (1991). Fractal and Chaotic Dynamics in Nervous Systems. Progress in Neurobiology, 36, pp. 279-308, 1991. http://citeseer.nj.nec.com/king91fractal.html

- Klein, S. (2002). Libet's Timing of Mental Events: Commentary on the Commentaries. Consciousness and Cognition. 11, 326-333. http://cornea.berkeley.edu/pubs/ccog_2002_0580-Klein-Commentary_on_the_commentaries.pdf

- Kleinschmidt, A., Bichel, C. Zeki, S. and Frackowiak, R.S.J. (1998). Human brain activity during spontaneously reversing perception of ambiguous figures. Proc R Soc Lond B Biol Sci. 1998 Dec 22;265(1413):2427-33

- Konishi, S., Wheeler, M.E., Donaldson, D.I., Buckner, R.L.. (2000) Neural correlates of episodic retrieval success. NeuroImage: 12: 276-286.

- Kornhuber HH, Deecke L: HirnpotentialÃ¤nderungen beim Menschen vor und nach WillkÃ¼rbewegungen, dargestellt mit Magnetbandspeicherung und RÃ¼ckwÃ¤rtsanalyse. PflÃ¼gers Arch. ges. Physiol. 281 (1964) 52.

- Kosslyn, S.M., Thompson, W.L. (2003). When is early visual cortex activated during visual mental imagery? Psychol. Bull. 129 (2003)723-746.

- Kreiman, G., Koch, C. and Fried, I. (2000). Imagery neurons in the human brain. Nature (2000) 408: 357-361 http://www.cns.caltech.edu/~gabriel/publications/imagination1.html

- Kreiman, G., Fried, I., & Koch, C. (2001). Single neuron responses in humans during binocular rivalry and flash suppression. Journal of Vision, 1(3), 131a,

http://journalofvision.org/1/3/131

- Kreiman, G., Fried, I. And Koch, C. (2002) Single Neuron responses in the human brain during flash suppression. Proceedings of the National Academy of Science USA, 99:8378-8383 http://www.ai.mit.edu/projects/cbcl/publications/ps/kreimanetal_flash-suppression.pdf

- Laureys, S. et al. (2002). Brain function in the vegetative state. Acta neurol. belg., 2002, 102, 177-185 http://www.jsmf.org/meetings/2003/nov/LaureysANB2002.pdf

- Lee. S, and Blake, R. (2002). V1 activity is reduced during binocular rivalry. Journal of Vision (2002) 2, 618-626 http://journalofvision.org/2/9/4/article.pdf

- Leopold, D.A. and Logothetis, N.K. (1999). Multistable phenomena: changing views in perception. Trends in Cognitive Sciences. 3: 254-264. http://www.kyb.tuebingen.mpg.de/publications/pdfs/pdf201.pdf

- Libet, B., Alberts, E.W., Wright, E. W., Jr., Feinstein, B.(1967). Responses of Human Somatosensory Cortex to Stimuli Below Threshold for Conscious Sensation. Science 158, 1597-1600.

- Libet, B., Wright, E. W., Jr., Feinstein, B., and Pearl, D.(1979). Subjective referral of the timing for a conscious sensory experience. Brain, 102, pp. 192-224.

  http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?holding=npg&cmd=Retrieve&db=PubMed&list_uids=427530&dopt=Abstract

- Linden, D.E.J. (2005) The P300: Where in the Brain Is It Produced and What Does It Tell Us? The Neuroscientist, Vol. 11, No. 6, 563-576 (2005) http://nro.sagepub.com/cgi/reprint/11/6/563.pdf

- Llinas, R., Ribary, U., Contreras, D. & Pedroarena, C. (1998). The neuronal basis for consciousness. Philos. Trans. R. Soc. London Ser. B.353, 1841-1849 http://big.cs.uiuc.edu/anncbt-papers/fall05/Llinas.the-neuronal-basis-of-consciousness.pdf

- Logothetis, N.K., Leopold, D.A. and Sheinberg, D.L. (1996). What is Rivalling During Binocular Rivalry. Nature, 380:621-624. http://www.nature.com/cgi-taf/DynaPage.taf?file=/nature/journal/v380/n6575/abs/380621a0.html

- Loveblad, KO. et al. (1999). Silent functional magnetic resonance imaging demonstrates focal activation in rapid eye movement sleep. Neurology 1999;53:2193Z

- Lumer, E.D., Friston, K.J. & Rees, G. (1998) Neural correlates of perceptual rivalry in the human brain. Science. 280(5371):1930-4.

- Macdonald, K.D., Fifkova, E., Jones, M.S. and Barth, D.S. (1998) Focal Stimulation of the Thalamic Reticular Nucleus Induces Focal Gamma Waves in Cortex. The Journal of Neurophysiology Vol. 79 No. 1 January 1998, pp. 474-477 http://jn.physiology.org/cgi/content/full/79/1/474

- Maia, T.V. & Cleeremans, A. (2005). Consciousness: converging insights from connectionist modeling and neuroscience. TRENDS in Cognitive Sciences Vol.9 No.8

August 2005 http://srsc.ulb.ac.be/axcWWW/papers/pdf/05-TICS.pdf

- Marcel AJ. (1998). Blindsight and shape perception: deficit of visual consciousness or visual function? Brain 1998, 121, 1565-1588. http://brain.oupjournals.org/cgi/reprint/121/8/1565.pdf

- Marois, R., Yi, D-J., & Chun, M.M. (2004). The Neural Fate of Consciously Perceived and Missed Events in the Attentional Blink. Neuron, 41, 3: 465-472.

- Mashour, G.A., Tononi, G. and Massimini, M. (2005). Cognitive Unbinding in Sleep and Anesthesia. Science 16 December 2005: 1768-1769.

- Moutoussis, K. and Zeki, S. (1997). A direct demonstration of perceptual asynchrony in vision. Proceedings of the Royal Society of London, Series B: Biological Sciences, 264, 393 - 399. http://juno.ingentaselect.com/vl=6492085/cl=26/nw=1/rpsv/cgi-bin/cgi?body=linker&ini=nlm&reqidx=issn=0962-8452vl=264is=1380yr=1997mn=Marpg=393

- Moutoussis, K. and Zeki, S. (2002). Responses of Spectrally Selective Cells in Macaque Area V2 to Wavelengths and Colors. J Neurophysiol 87: 2104-2112, 2002. http://jn.physiology.org/cgi/content/full/87/4/2104

- Moutoussis, K. and Zeki, S. (2002). The relationship between cortical activation and perception investigated with invisible stimuli. Proc. Nat. Acad. Sci. USA, 99(14):9527-9532. http://www.pnas.org/cgi/content/full/99/14/9527

- Muckli, L., Singer, W., Zanella, F.E. and Goebel, R. (2002) Integration of multiple motion vectors over space: an fMRI study of transparent motion perception. Neuroimage. 2002 Aug;16(4):843-56.

- Munhall, K.G. (2001). Functional Imaging During Speech Production. Acta Psychologica 107 (2001) 95-117. http://pavlov.psyc.queensu.ca/~munhallk/07_acta_version.pdf

- Munk MHJ, Roelfsema PR, KÃ¶nig P, Engel AK, and Singer W. (1996). Role of reticular activation in the modulation of intracortical synchronization Science. 272, 271-274.

- Newman, J. & Baars, B.J. (1993) A neural attentional model for access to consciousness: A Global Workspace perspective. Concepts in Neuroscience, 4:2, 255-290 http://cogprints.ecs.soton.ac.uk/archive/00000073/

- Nieuwenhuis S, Aston-Jones G & Cohen JD (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. Psychological Bulletin, 131(4), 510-532. http://www.csbmb.princeton.edu/conte/pdfs/project5/Proj5Pub06.pdf

- Pantev C (1995) Evoked and induced gamma-band activity of the human cortex. Brain Topogr 7:321-330

- Pasley BN, Mayes LC, Schultz RT (2004) Subcortical discrimination of unperceived objects during binocular rivalry. Neuron 2004, 42:163-172.

- Pessoa, L. (2005) To what extent are emotional visual stimuli processed without attention

and awareness? Current Opinion in Neurobiology 2005, 15:188-196 http://www.brown.edu/Research/LCE/publications_files/Pessoa_CONB_2005.pdf

- Pettigrew, J.D. (2001) Searching for the Switch: Neural Bases for Perceptual Rivalry Alternations. Brain and Mind 2: 85-118, 2001. http://www.uq.edu.au/nuq/jack/searchswitch.pdf

- Raymond, J. E, Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? Journal of Experimental Psychology. Human Perception and Performance, 18 (3), 849-60.

- Rees, G., Kreiman, G., & Koch, C. (2002) Neural Correlates of Consciousness in Humans. Nature Reviews Neuroscience 3, 261 -270 (2002) http://www.nature.com/cgi-taf/DynaPage.taf?file=/nrn/journal/v3/n4/abs/nrn783_fs.html&dynoptions=doi1069842948

- Rensink, RA. (2000). When Good Observers Go Bad: Change Blindness, Inattentional Blindness, and Visual Experience. PSYCHE, 6(09), August 2000. http://psyche.cs.monash.edu.au/v6/psyche-6-09-rensink.html

- Roy, J.E. & Prichep, L.S. (2005) The Anesthetic Cascade: A Theory of How Anesthesia Suppresses Consciousness Anesthesiology: Volume 102(2) February 2005 pp 447-471 http://www.anesthesiology.org/pt/re/anes/fulltext.00000542-200502000-00030.htm

- Schiff, N.D., & Plum F. 1999. The Neurology of Impaired Consciousness: Global Disorders and Implied Models. [Target article]. Association for the Scientific Study of Consciousness. http://www.phil.vt.edu/assc/niko.html

- Schiff ND, Plum F. (2000) The Role of Arousal and "Gating" Systems in the Neurology of Impaired Consciousness. Journal of Neurophysiology. 17(5):438-452. (http://www.uni-tuebingen.de/cog/lehre/ss2003/seminar_bewusstsein/epapers/SchiffPlum00.pdf

- Sheinberg, D.L. and Logothetis, N.K. (1997). The role of temporal cortical areas in perceptual organization. Proc. Natl. Acad. Sci. USA 94, 3408-3413. http://www.pnas.org/cgi/content/full/94/7/3408

- Simons, D.J. & Rensink, RA. (2005). Change blindness: Past, present, and future. Trends in Cognitive Sciences, Vol.9, pp16-20, No.1 January 2005. http://www.psych.ubc.ca/~rensink/publications/download/S&R-TICS-05a.pdf

- Singer, W. & Gray, C.W. (1995). Visual feature integration and the temporal correlation hypothesis. Annu. Rev. Neurosci., 18, pp. 555-86, 1995.

- Singer, W. (2001). Consciousness and the Binding Problem. Annals of the New York Academy of Sciences 929:123-146 (2001)

- Solms, Mark (2000) DREAMING AND REM SLEEP ARE CONTROLLED BY DIFFERENT BRAIN MECHANISMS, Behavioral and Brain Sciences 23 (6): 793. http://www.bbsonline.org/Preprints/OldArchive/bbs.solms.html

- Sukov, W. & Barth, D.S. (2001). Cellular Mechanisms of Thalamically Evoked Gamma Oscillations in Auditory Cortex. J Neurophysiol 85: 1235-1245, 2001;

http://jn.physiology.org/cgi/content/full/85/3/1235

- Takeuchi, T., Miyasita, A., Inugami, M. & Yamamoto, Y. (2001). Intrinsic dreams are not produced without REM sleep mechanisms: evidence through elicitation of sleep onset REM periods. Journal of Sleep Research Volume 10 Issue 1 Page 43 - March 2001 http://www.blackwell-synergy.com/links/doi/10.1046/j.1365-2869.2001.00237.x/full/

- TamÃ¡s, G., Buh, E.H., LÃ¶rincz, A., and Somogyi, P. (2000). Proximally targeted GABAergic synapses and gap junctions synchronize cortical interneurons. Nature Neuroscience volume 3 no 4 april 2000. http://mrcanu.pharm.ox.ac.uk/pdfs/TamasNatureNeurosci2000.pdf

- Tanaka, Y. & Sagi, D. (1998). A perceptual memory for low contrast visual signals. Proc. Natl. Acad. Sci. USA Vol 95:12729-12733. http://www.weizmann.ac.il/~masagi/TS98b.pdf

- Tong, F. (2001). Competing Theories of Binocular Rivalry: A Possible Resolution. Brain and Mind 2: 55-83, 2001 http://www.princeton.edu/~ftong/Tong_Brain&Mind2001.pdf

- Tong, F. (2002). Primary Visual Cortex and Visual Awareness. Nature Reviews Neuroscience 4, 219 (2003) http://www.princeton.edu/~ftong/Tong_NRN2003.pdf

- Tootell, R.B., Reppas, J.B., Dale, A.M., Look, R.B., Sereno, M.I., Malach, R., Brady, T.J. and Rosen, B.R. (1995) Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. Nature. 1995 May 11;375(6527):139-41.

- Tononi, G., Srinivasan, R., Russell, P.D. & Edelman, G.M. (1998). Investigating neural correlates of conscious perception by frequency-tagged neuromagnetic responses . Proc. Natl. Acad. Sci. USA. Vol. 95, Issue 6, 3198-3203, March 17, 1998 http://www.pnas.org/cgi/content/full/95/6/3198

- Thornton, I.M., Pinto, J. and Shiffrar, M. (1998). THE VISUAL PERCEPTION OF HUMAN LOCOMOTION. COGNITIVE NEUROPSYCHOLOGY, 1998, 15 (6/7/8), 535-552. http://psychology.rutgers.edu/~mag/TPS98.pdf

- Traub, D.R., Whittington, M.A., Buhl, E.H., Jefferys, J.G.R., & Faulkner, H.J. (1999). On the Mechanism of the Frequency Shift in Neuronal Oscillations Induced in Rat Hippocampal Slices by Tetanic Stimulation. The Journal of Neuroscience, February 1, 1999, 19(3):1088-1105 http://www.jneurosci.org/cgi/content/full/19/3/1088

- Traub, R.D. (2003). Fast Oscillations and Epilepsy. Epilepsy Curr. 2003 May; 3(3): 77-79. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=321178

- Tse, P.U., Susana Martinez-Conde, S., Schlegel, A.A. and Macknik, S.L. (2005). Visibility, visual awareness, and visual masking of simple unattended targets are confined to areas in the occipital cortex beyond human V1/V2. Proc. Natl. Acad. Sci. USA November 22, 2005 vol. 102 no. 47. 17178-17183. http://www.pnas.org/cgi/content/full/102/47/17178

- Vogel, E.K., Luck, S.J. & Shapiro, K.L. (1998) Electrophysiological Evidence for a Postperceptual Locus of Suppression During the Attentional Blink. Journal of Experimental Psychology: Human Perception and Performance 1998, Vol. 24. No. 6,1656-1674 http://www.psychology.uiowa.edu/faculty/Luck/lucklab/pdfs/Vogel_1998_JEPHPP.pdf

- Wehrle, R., Czisch, M., Kaufmann, C., Wetter, T.C., Holsboer, F., Auer, D.P., and Pollmacher, T. (2005). Rapid eye movement-related brain activation in human sleep: a functional magnetic resonance imaging study. NeuroReport 16:853-857 http://www.univie.ac.at/mcogneu/lit/wehrle.pdf

- White, NS., Alkire, MT. (2003). Impaired thalamocortical connectivity in humans during general-anesthetic-induced unconsciousness. NeuroImage. Volume 19, Issue 2 , June 2003, Pages 402-411. http://smallurl.com/?i=2597 (http://www.sciencedirect.com/science/article/B6WNP-48GP5CG-5/1/cedb33dc0bec4301d1960c7c27cea70f)

- Wheeler, M.E., Buckner, R.L.. (2003) Functional dissociation among components of remembering: control, perceived oldness, and content. J. Neurosci. 23, 3869-3880.

- Williams LM, Liddell BJ, Rathjen J, Brown KJ, Shevrin H, Gray JA, Phillips M, Young A & Gordon E (2004). Mapping the time course of nonconscious and conscious perception of fear: An integration of central and peripheral measures. Human Brain Mapping,21, 64 - 74 http://psychmed.iop.kcl.ac.uk/neuroscience-and-emotion/downloads/46.pdf

- Wunderlich, K. Schneider, K.A., & Kastner, S. (2005).Neural correlates of binocular rivalry in the human lateral geniculate nucleus. Nature Neuroscience 8, 1595 - 1602 (2005)

- Yamamoto, T. & Katayama, Y. (2005). Deep brain stimulation therapy for the vegetative state. in Neuropsychol Rehabil. 2005 Jul-Sep ; 15(3-4): 406-13

- Zeki, S., & Bartel, A. (1999). Toward a Theory of Visual Consciousness. Consciousness & Cognition, 8, 225-259.

Behaviourism and Consciousness

References

- Behaviourism. Graham, G. (2002) Stanford Encyclopedia of Philosophy
- Watson, J.B. (1907) Studying the Mind of Animals The World Today, 12, 421-426.
- Watson, J.B. (1913) Psychology as the Behaviorist Views it. Psychological Review, 20, 158-177
- Lektorsky, V.A. (1980)Subject Object Cognition.
- Vygotsky L.S. (1925) 'Consciousness as a problem in the psychology of behavior' Undiscovered Vygotsky: Etudes on the pre-history of cultural-historical psychology (European Studies in the History of Science and Ideas. Vol. 8), pp. 251-281
- Gibson, J. J. (1950). The perception of the visual world. Boston: Houghton-Mifflin
- Gibson, J. J.(1979). The ecological approach to visual perception. Boston: Houghton-Mifflin
- Ryle, G. (1949). The concept of mind. Harmondsworth, Middlesex: Penguin.

[edit]

Models of Access Consciousness

There have been numerous attempts to model reflex and access consciousness. These models, being connectionist and information systems based, do not model phenomenal consciousness but are essential steps in understanding global brain function.

Neural networks

Neural networks achieve information processing by establishing connections between processing units in a system of processors that have similar characteristics. Neural networks are used for classifying data. The processing units serve the function of both filtering and storing information.

This is a stub and requires expansion

Classification of sensory stimuli

The path from transducers to a single neuron that responds to a single complex stimulus.

This is a stub and requires expansion

Classification of motor control

From premotor activity to skilled behaviour.

This is a stub and requires expansion

Olfaction: classification out of chaos?

This is a stub and requires expansion

Quantum information processing

*This is a **stub** and needs expanding*

Part IV: Explanations of Consciousness

Contemporary Explanations of Consciousness

This section is about the types of theory that have been advanced to explain consciousness. Specific explanations should be entered as separate pages.

Introduction

Explanations of consciousness fall into four broad categories, those that attempt to explain the empirical experience called consciousness with scientific theories, those that seek to find some way in which consciousness could be explained by digital computers or nineteenth century materialism by redefining or eliminating experience and those that regard consciousness as inexplicable or supernatural.

Identity theory of mind

The identity theory of mind, or type physicalism, holds that the mind is identical to the brain. Type physicalists identify qualia and the form of experience with brain activity. They argue that "mind states" have physical causes and physical effects - thus the mind states themselves must be physical; a non-physical "middle step" is superfluous.

Type physicalism has not yet gained widespread support because although brain activity that correlates with experience has been found everywhere in the brain, no set of brain activity that is phenomenal consciousness itself has yet been found - although this is not surprising because neuronal spike activity is unlikely to host phenomenal consciousness - see scientific theories of consciousness.

Functionalism

Functionalism was developed as a theory of the mind-body problem because of objections to identity theory and logical behaviourism. Its core idea is that the mental states can be accounted for without taking into account the underlying physical medium (the neurons), instead attending to higher-level functions such as beliefs, desires, and emotions. It is a theory of behaviour and access consciousness and so from the outset avoids any explanation of phenomenal consciousness, substituting beliefs and judgements (functions) for entities such as qualia.

According to functionalism, the mental states that make up consciousness can essentially be defined as complex interactions between different functional processes. Because these processes are not limited to a particular physical state or physical medium, they can be realized in multiple ways, including, theoretically, within non-biological systems.This affords consciousness the opportunity to exist in non-human minds that are based on algorithmic processors such as digital computers. This is a highly contentious conjecture although non-functionalist physicalists might agree that machines that are not digital computers could possess consciousness through an identity theory of mind - see The problem of machine and digital consciousness.

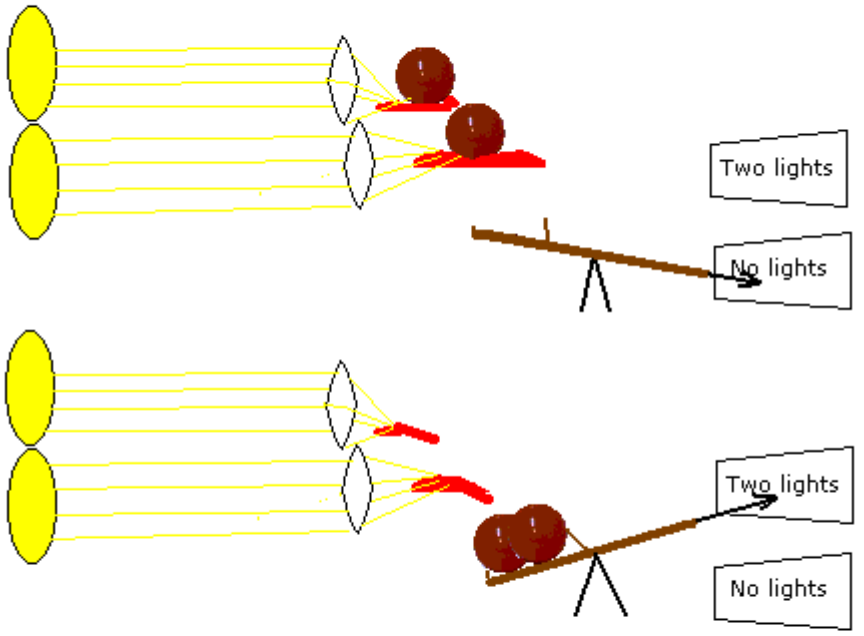
Functionalism's explanation of consciousness, or the mental, is best understood when considering the analogy made by functionalists between the mind and the modern digital computer. More specifically, the analogy is made to a "machine" capable of computing any given algorithm (i.e. a Turing machine). This machine would involve:

Data input (the senses in humans), data output (both behaviour and memory), functional states (mental states), the ability to move from one functional state into another, and the definition of functional states with reference to the part they play in the operation of the entire entity - i.e. in reference to the other functional states.

So long as the same process was achieved, the "physical stuff" -- that being computer hardware or biological structure -- could achieve consciousness. This combination of data input, data output, functional states and movement from state to state is shown in the model system in the illustration below.

Functionalism

According to functionalism a stimulus response system is conscious. In the model below the heat from lights melts some wax which allows two falling balls to provide the output "two lights".



Which part of the system observes the event? Would the addition of another lever and another set of balls with a pointer to a sign saying "I am observing the event" truly give rise to observation? Or does the system only make sense because you are a conscious observer with an active imagination?

Many naive proponents of functionalism do not realise that digital computers are conceptually equivalent to very large systems of balls and levers.

This variety of functionalism was developed by Hilary Putnam. One of the major proponents of functionalism is Jerry Fodor.

Further reading:

Block, N. (1996). The Encyclopedia of Philosophy Supplement, Macmillan, 1996 http://www.nyu.edu/gsas/dept/philo/faculty/block/papers/functionalism.pdf

Dualism

Substance dualism

This theory proposes that phenomenal experience occurs in a non-physical place. In Cartesian Dualism the non-physical place is a point-soul that looks out at the brain. In Reid's Natural Dualism the non-physical place is a point-soul that looks out at the world.

Property dualism

Property dualism asserts that when matter is organized in the appropriate way (i.e., organized in the way that living human bodies are organized), mental properties emerge. Property dualism is a branch of emergent materialism. The appeal to emergentism deserves closer attention. Scientific theories often deal with emergent phenomena, for instance an enzyme consists of carbon, hydrogen, nitrogen, manganese and oxygen and from this catalytic action emerges. The theory of enzyme structure and the action of this structure on the substrate explains how this emergence occurs. Notice that the theory of enzymes explains the emergence of catalytic activity; emergence does not explain the theory. In science the statement that some property will 'emerge' means that there will be a theory that accounts for this property. Property dualism, by appealing to emergence, is stating that some theory of consciousness will be possible. In other words it is an explanation that proposes that the explanation is yet to be known.

Intentionalism

[edit]

Higher order thought

This section is a stub and needs expansion

[edit]

Eliminativism

Eliminative materialism is the school of thought that argues for an absolute version of materialism and physicalism with respect to mental entities and mental vocabulary. It principally argues that our common-sense understanding of the mind (described with the pejorative term 'folk psychology' by eliminativists) is not a viable theory on which to base scientific investigation, and therefore no coherent neural basis will be found for many such everyday psychological concepts (such as belief or intention) and that behaviour and experience can only be adequately explained on the biological level.

Eliminative materialists therefore believe that consciousness does not exist except as an epiphenomenon of brain function and some believe that the concept will eventually be

eliminated as neuroscience progresses. Similarly, they argue that folk psychological concepts such as belief, desire and intention are illusory and therefore do not have any consistent neurological substrate.

Proponents of this view often make parallels to previous scientific theories which have been eliminated, such as the four humours theory of medicine, the phlogiston theory of combustion and 'vital force' theory of life. In these cases, science has not produced more detailed versions of these theories, but rejected them as obsolete. Eliminative materialists argue that folk psychology is headed the same way. According to W.V. Quine it will take tens of years before folk psychology will be replaced with real science. Eliminativism is novel however because it uses theory based on nineteenth century materialism to reclassify all observation as theory. This use of theory to eliminate observation is highly unusual and suspect (see Phenomenal consciousness and access consciousness).

Eliminative materialism was first defended by W.V. Quine, Paul Feyerabend, and Richard Rorty. This view is most associated with philosophers Paul and Patricia Churchland, although philosophers such as Daniel Dennett would also consider themselves eliminativists for many aspects of psychology. Philosopher Dale Jacquette has claimed that Occam's Razor is the rationale behind eliminativism and reductionism.

The most common argument against eliminative materialism the argument from qualia, which is deployed in various forms by Thomas Nagel, Frank Jackson, and many others. Perhaps the most powerful argument against eliminativism is that experience itself is many things simultaneously; it is, as Aristotle points out, immediately things in space and hence is not composed of judgements.

Mysterianism

New Mysterianism is a philosophy proposing that certain problems (in particular, consciousness) will never be explained.

Owen Flanagan noted in his 1991 book "Science of the Mind" that some modern thinkers have suggested that consciousness might never be completely explained. Flanagan called them "the new mysterians" after the rock group ? and the Mysterians. The term originated with the Japanese alien-invasion film The Mysterians. The "old mysterians" are thinkers throughout history who have put forward a similar position. They include Leibniz, Dr Johnson, and Thomas Huxley. The latter said, "How is it that anything so remarkable as a state of consciousness comes about as a result of irritating nervous tissue, is just as unaccountable as the appearance of the Djin, when Aladdin rubbed his lamp." [6, p. 229, quote]

Noam Chomsky distinguishes between problems, which seem solvable, at least in principle, through scientific methods, and mysteries which do not, even in principle. He notes that the cognitive capabilities of all organisms are limited by biology, e.g. a mouse will never speak. In the same way, certain problems may be beyond our understanding.

The term New Mysterianism has been extended by some writers to encompass the wider philosophical position that humans don't have the intellectual ability to understand many hard problems, not just the problem of consciousness, at a scientific level. This position is also known as Anti-Constructive Naturalism.

For example, in the mind-body problem, emergent materialism claims that humans aren't smart enough to determine "the relationship between mind and matter." [4] Strong agnosticism is a religious application of this position.

Colin McGinn is the leading proponent of the New Mysterian position.

Critics argue this philosophy isn't useful and encourages capitulation. One critic noted:

the extreme "Mysterian" position, that there are vital issues forever beyond our reach, is in many ways deeply unsatisfying. [7]

References

[1] McGinn, Colin - The Problem of Consciousness

[2] McGinn, Colin - Problems in Philosophy: the limits of enquiry

[3] McGinn, Colin - The Mysterious Flame

[4] Blackburn, Simon - Think: A compelling introduction to philosophy, chapter two

[5] Flanagan, Owen - The Science of the Mind (1991) 2ed MIT Press, Cambridge

[6] Horgan, John - The Undiscovered Mind (1999), Phoenix, ISBN 0753810980

Idealism and panpsychism

Idealism

[edit]

Panpsychism

If you are here, reading about panpsychism, then it may be because you still have no clue as to what consciousness is. Or, you have read philosophers say ' Without consciousness the Mind-Body problem would be much less interesting, with consciousness it seems hopeless' T. Nagel, ' What it's Like to be a Bat'. You may have reached this section rationally, logically. It follows the previous section, you are reading linearly. Basically you are thinking and acting scientifically/ philosophically. Or, you may have read " Panpsychism " and thought ' strange, flaky, interesting '. You are following your feelings, and close behind are your intuitions. This kind of thinking could be called Magical thought. There is more than one way to be aware of the world. Although a panpsychic might say consciousness is everything, it is not every where the same thing. It has an infinite number of aspects, forms and patterns. We arrive at panpsychism intuitively, by following our feelings. Science and philosophy represent one kind of awareness, one aspect of consciousness.Magical thinking is another aspect of consciousness. " ...intuitions are important and that even if expressed vaguely they can serve as useful pointers to those seeking a more complete account of the mind" David Skrbina, ' Panpsychism in the West ' p.3.

Put 2 Greek god's together and you end up with a term that appears more mythological and

intuitive than scientific or philosophical. Pan- Psyche (ism) ! That is the place we begin to understand what the term means. The literature and art of the Ancient Greeks express the mythical and intuitive perspective. ' Modern theories of panpsychism have their roots in the mythology and spiritualism of the pre-classical world. ' David Skrbina, 'Panpsychism in the West ' p.23 (MIT press). That perspective begins to change with the emergence of science and philosophy or, correctly, science and philosophy emerge when the perspective changes. Greek consciousness changed. (see the section on Homer above) ' They transformed the mythological, pre-historic animist worldview into rational and logical theories of the cosmos' (ibid).

Scientific Theories of Consciousness

Theories of Consciousness

Some recent scientific theories of consciousness are tabulated below. The extent to which they account for the phenomenon of consciousness is shown.

It is remarkable that many of the theories are consistent with one another. As in the tale of the 'blind men and the elephant' some of the theories seem to describe the trunk, some the tail etc. but they all seem to be part of the same elephant! The convergence of the theories is shown in the illustration below:

A Summary of some of the theories of consciousness

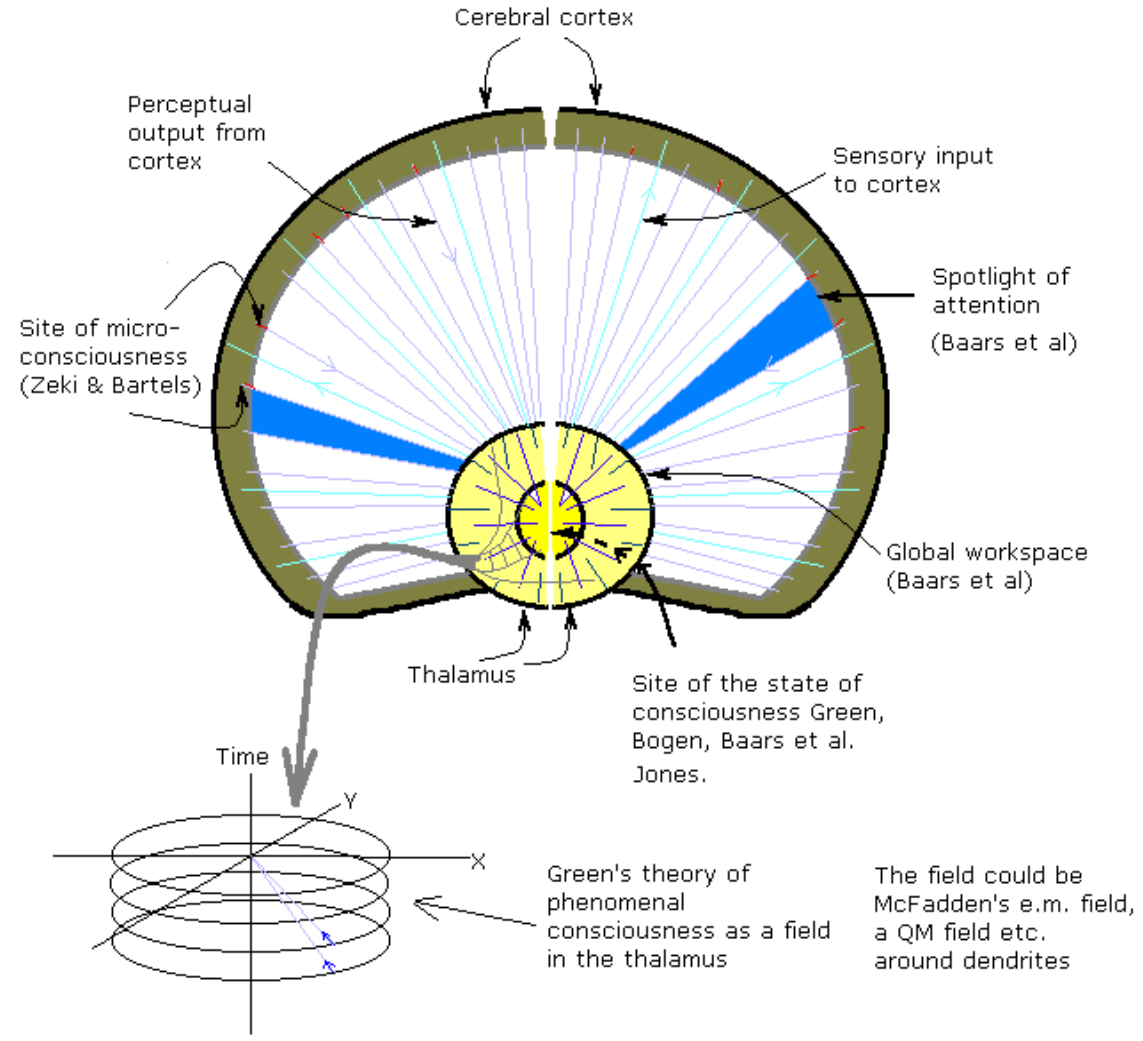


Table of theories

A = Model of observer's view

B = Model of Anaesthetic Action in thalamus

C = Explanation of Libet's data

D = Explanation of unconscious but active cerebral cortex

E = Explanation of knowing you know

F = Explanation of non-computability

G = Binding (simultan of relevant data)

H = Extended present

I = Quantum state vecto

Name	Author/Ref	A	B	C	D	E	F	G	H	I
Microconsciousness	Zeki, S., & Bartel, A. (1999) Toward a Theory of Visual Consciousness. Consciousness & Cognition, 8, 225-259.	N	N	N	N	N	N	Y	N	N
Geometrical Phenomenalism	Green, A. (2003) Geometrical phenomenalism	Y	Y	Y	Y	Y	?	Y	Y	Y
ORCH-R	Hameroff, S & Penrose, R. 1989	N	N	N	N	N	Y	Y	N	Y
Quantum Brain Model	Ricciardi, L. M. and H. Umezawa, 1967. Brain physics and many-body problems, Kibernetik 4, 44-48. http://arXiv.org/abs/q-bio/0309009	N	N	N	N	N	?	Y	N	Y
Many Minds	Donald, M. 1990. Quantum Theory and the Brain. Proc R Soc Lond. A427 43-93. http://xxx.lanl.gov/PS_cache/quant-ph/pdf/9904/9904001.pdf	Y	N	N	N	?	?	Y	Y	Y

A = Model of observer's view

B = Model of Anaesthetic Action in thalamus

C = Explanation of Libet's data

D = Explanation of unconscious but active cerebral cortex

E = Explanation of knowing you know

F = Explanation of non-computability

G = Binding (simultan of relevant data)

H = Extended present

I = Quantum state vecto

Name	**Author/Ref**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
Dual-Time Supercausality	King, C.C. 1989. Physics Essays 2/2 128-151. http://www.math.auckland.ac.nz/~king/Preprints/Transup.htm	N	N	N	N	Y	Y	Y	?	Y
Spin Mediated Consciousness	Hu, H. & Wu, M. 2002. Spin-Mediated Consciousness Theory: An Approach Based On Pan-Protopsychism. http://cogprints.ecs.soton.ac.uk/archive/00002579/	N	N	N	N	N	?	Y	N	Y
Quantum Theory of Consciousness (synaptic cleft)	Walker, E.W. 1998. the *Noetic Journal*, *1*, 100-107, 1998 http://users.erols.com/wcri/CONSCIOUSNESS.html	N	N	N	N	N	N	Y	N	Y

A = Model of observer's view

B = Model of Anaesthetic Action in thalamus

C = Explanation of Libet's data

D = Explanation of unconscious but active cerebral cortex

E = Explanation of knowing you know

F = Explanation of non-computability

G = Binding (simultan of relevant data)

H = Extended present

I = Quantum state vecto

Name	Author/Ref	A	B	C	D	E	F	G	H	I
Global Workspace Theory	Baars, B. 1988. A cognitive theory of consciousness. Cambridge University Press, New York http://www.ceptualinstitute.com/genre/baars/baarsBrain.htm	Y	Y	Y	Y	N	N	Y	N	N
Topological Geometrodynamics (TGD) Inspired Theory of Consciousness	PitkÃ¤nen, M. 199?. Topological Geometrodynamics http://www.physics.helsinki.fi/~matpitka/mainpage.html	Y	?	N	N	Y	Y	Y	Y	Y
The Conscious Electromagnetic Field Theory	McFadden, J.J. 2002 http://www.surrey.ac.uk/qe/cemi.htm	Y	?	N	N	N	N	Y	N	?
Name	**Author/Ref**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
Real Time Consciousness	Smythies, J. 2003.Journal of Consciousness Studies 10:3 47-56 http://www.imprint.co.uk/pdf/smythies.pdf	Y	N	N	N	Y	?	Y	Y	?
Consciousness as memory	Gerald Edelman's theory Proc Natl Acad Sci U S A. 2003 April 29; 100(9): 5520â€“5524	N	N	N	N	N	N	N	N	N

Other Explanations

Conscousness only

Consciousness-only is the foundation of a buddhist theory known as vijnanavada. Proponents suggest that the sum of experience exists only in our minds. Philosophers recognize this view as subjective idealism. Consciousness-only views can also be found in taoist philosophy, notably Lao Tzu and Chuang Tzu.

Bps model

This brief summary is an illustration of an attempt to model a multidisciplinary biopsychosocial (bps) understanding of self-consciousness seen from the perspective of both scientific methodology and metaphysical logic where the empirical and the inferential provide a seamless blend of the ontological brain with the epistemological mind.

The achievement of self consciousness is the crucial mental state allowing the human species to monitor the equilibrium state of biopsychosocial ongoing contingencies especially when confronting life-threatening circumstances. The inherited proto-semantics and acquired language guide the required recursive co-generation of the appropriate language and thought to meet the contingency. Thus informed, it allows humans to elaborate effective adaptive short and long range responses.

Definition of terms

Bps model uses some unusual definitions of terms. These are explained below.

"Sense-phenomenal awareness" is defined as an unconscious, life-preserving, adaptive reflex response which may occur without qualia. It originates at a sensory receptor, wherever located in the body economy, and ends at an effector organ, glandular or muscular. - Phenomenal consciousness/awareness is a term normally reserved for experience containing qualia in other analyses.

System/network "awareness" is defined in the bps model as that unconscious processing occuring during the integration of the participating neural network modules leading to a stereotyped adaptive response. - normally awareness is defined as knowledge that a conscious state is present.

Sense-phenomenal awareness may become a conscious experience when relevant inferential networks (e.g., memory, emotions, etc.)are subsequently accessed, including inner-language processors. When experiences are recalled the qualia that arise are called "conceptual qualia".

"Access consciousness" is described as being initially an unconscious process that makes it possible for a life-preserving, reflex-driven and 'unconscious' sense-phenomenal state of mind to become conscious by making use of available, pertinent and concurrent mental states to interact with the novel sense-phenomenal input, a potentially life-threatening event.

"Proto-linguistic organ" or 'plo' is described as the first line of defense to guard against life-threatening stimuli arising from sense-phenomenal inputs (external, visceral or propioceptive). Housed in the amygdaloidal complex, it represents the inherited proto-semantic (primitive 'meanings') database responsible for activating the corresponding unconscious fight/flight adaptive Cannon effector response.

"Proto-semantic" input from plo is described as a required initial participant in the subsequent recursive co-generation of inner language and thought as may be required in the eventual elaboration of "conceptual consciousness".

Higher order consciousness theory

The 'bps' model of 'consciousness' is a high order consciousness theory in which an unconscious, non inferential phenomenal state (established from either online sensory receptor input or offline memory input), when confronting a novel life-threatening event, triggers an initially unconscious access intermediate stage where relevant modular networks are incorporated including Broca's language processor recursively co-generating in the process the 'inner language' narrative state and accompanying thought, a conscious high order mental state, all of which causally precedes (or is simultaneous with) the adaptive response (if any, as we see in dreams).

Notice that bps considers phenomenal states to be non-conscious, this would confuse the ordinary reader who expects the Kantian term "phenomenal" to be equivalent to the term "conscious experience". Only the higher order mental state is regarded as "conscious".

The 'bps' model basically describes two co-existing, ongoing mental states, one non-inferential subconscious 'gut feeling' inner sense (BOP, a variant of Lycan's 1996 HOP) and an initially non-inferential unconscious accessing of narrative pathways leading to (recursive co-generation of 'inner language' and thought is an open option) the eventual production of higher order thought (HOT) whose content is the feeling that oneself is the subject of self-consciousness.

In other words, according to the 'bps' theory, feelings are not part of consciousness until higher order thought occurs, ie, qualia needs a context.

In 'bps' theory not even self-consciousness, of which 'qualia' may arguably be considered a subset of, has revealed its constitutive secrets. This means that bps is a theory of brain processing rather than a theory of the content of consciousness (qualia) or consciousness itself except when it ventures into the postulate that language and self-consciousness are recursively co-generated or co-causal. More controversial is the mediation of the amygdaloid complex (plo) in providing inherited primitive 'meanings' (protosemantic codelets) to initiate Chomskian language processing and thought co-generation, i.e., protosemantics precedes syntax structuring. For a more complete exposition see: http://home.earthlink.net/~dr.ds/neurophilosophyofconsciousnesssummary/id1.html

Further Reading: http://delasierra-sheffer.net/

License

GNU Free Documentation License

Version 1.2, November 2002

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title,

year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties--for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements."

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole

aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

External links

- GNU Free Documentation License (Wikipedia article on the license)
- Official GNU FDL webpage

Retrieved from "http://en.wikibooks.org/wiki/Consciousness_studies/Print_version"