

# *On the Conceivability of Zombies*

By BRENT SILBY

Department Of Philosophy, University of Canterbury, New Zealand

Copyright (c) Brent Silby 1998

[www.def-logic.com/articles](http://www.def-logic.com/articles)

## **Introduction**

Consciousness lies at the very heart of the mind/body problem. So far it has been resistant to scientific analysis and explanation. Essentially, the mind/body problem consists in the realisation that while it may be possible to discover physical principles by which the mind works, the phenomenon of consciousness may escape the physical story. It seems that we have what the literature refers to as an *explanatory gap*, and that consciousness, by its very nature, cannot be accounted for by physical theories of the mind/brain. This type of thinking suggests a metaphysical dualism between the body and the mind. Anti-physicalist arguments have taken a similar form since Descartes, and often involve conceivability arguments. In other words, the anti-physicalist puts forward thought experiments that involve the conceptual possibility of certain scenarios (Levine 1993: p543). These thought experiments involve imagining situations in which a complete physical knowledge of the world (including the brain) fails to provide an account of conscious experience. For example, it is claimed that we can imagine a situation in which there exist creatures who are physically and behaviourally identical to us, yet who have no conscious experience. The conceivability of such creatures is supposed to show that consciousness is not logically supervenient on the physical.

It is something of an irony that while our minds are the one thing of which we are directly aware, the existence and nature of our minds is one of the most puzzling mysteries that we have ever grappled with. Being a part of the physical world, we would like to be able to explain the workings of our minds in scientific terms and there has been a great deal of success in this area. Neuroscientists are working hard to provide us with neural models of the brain and progress is being made in understanding the way the brain works. However, despite the progress being made, philosophers such as David Chalmers believe that *consciousness* escapes reductive explanations.

In his book 'The Conscious Mind', Chalmers makes significant use of thought experiments involving zombies. These thought experiments are supposed to show that consciousness is not *logically supervenient* on the

physical and that conscious states, or the properties of those states, called qualia, are resistant to physical and functional analysis.

In this paper I will first explain the concept of supervenience, then describe the problem of qualia and explain why qualia are a problem for physical theories of mind. I will then describe the zombie argument and will explain Chalmers' attempts to use it as an argument for the conclusion that qualia do not logically supervene on the physical. My goal will be to show that the supposed conceivability of zombies does not offer any serious refutation of physical or functional theories of the mind.

## 1: Supervenience

The notion of supervenience is an important part of Chalmers' argument. The concept of supervenience allows us to formally locate relations of dependence between sets of properties. If A-properties supervene on B-properties, then whenever you find B-properties, you will find A-properties. Now, there is an important distinction to be made between two types of supervenience - *logical* supervenience and *natural* supervenience. When we say that B-facts are *naturally* supervenient on A-facts, we are making a claim about the empirical world. For example, in our world the facts and laws of gravity supervene on the facts and laws relating to the mass of objects. There is a law-like correlation between how massive objects are and the gravitational forces that they give rise to. So, in this world (or universe), whatever object we choose to examine, that object will be surrounded by a gravitational field that is proportional to its total mass. There are no exceptions to this law. Gravity *naturally* supervenes on mass. However, we can imagine a possible universe in which the laws of physics are vastly different to the laws in our universe. Perhaps in our imagined universe there is no gravity. In our imagined universe gravity would not supervene on mass because in that universe gravity does not exist. Our ability to imagine such a universe shows us that gravity is not *logically* supervenient on mass. For B-properties to be *logically* supervenient on A-properties there must be no logically possible situations that share identical A-properties and yet remain distinct with regard their B-properties (Chalmers 1996: p35). Logical supervenience is defined in terms of logically possible worlds. For example, we could say that biological properties supervene logically on physical properties, because it would be logically impossible for a world to exist that was physically identical to ours and yet was biologically distinct (Chalmers 1996: p35). When all the physical properties of our world are in place, the biological properties come along at no extra charge. Even God could not create a universe that was physically identical to ours (including the position of every particle), while keeping it biologically distinct. If penguins exist in our world, then any world that is physically identical to ours also contains penguins. Biology is logically supervenient on the physical.

Now, Chalmers has made the claim that consciousness is not logically supervenient on the physical. So, for Chalmers, there is a possible world that is identical to our world in all physical respects, but is distinct in the fact that consciousness (in humans) does not exist. But what is Chalmers referring to when he speaks of consciousness? The word 'consciousness' can be used to refer to a number of different aspects of human life, but Chalmers has a specific aspect of consciousness in mind. He is concerned with the existence of qualia.

## 2: Qualia

One of the central problems in the contemporary philosophy of mind is the problem of qualia. The word 'qualia' refers to the intrinsic properties of our experience, including phenomenal properties such as colours, sounds and sensations of pains. These 'raw feels' seem to be intrinsic properties of our experience and yet cannot be found in the non-mental world. According to tradition, qualia are defined as ineffable, intrinsic, private and directly apprehensible in consciousness (Dennett 1988: p622). Sensations of colours, pains and sounds seem very real to us and yet when we turn to science for an explanation of the colour *blue*, all we learn is that there are certain wavelengths of electromagnetic radiation that are being reflected from the surfaces of various objects. There is no blue, just radiation. The question we are left with is: where is the *blueness* if it is not in the outside world? An intuitive answer to this question could go like this:

*When light enters our eyes, various frequencies of radiation cause the retina's cones to activate. This initiates an electrochemical cascade by which signals are sent to the visual cortex and a neural state is set up. The visual cortex is connected to a 'box' that is filled with 'an instance of' the colour, which can then be examined by other parts of the brain in order to produce behavioural output.*

This story answers the question badly. It may *feel* as if the blue colour is there, in experience, but seeming like something is the case does not show that it is the case. Daniel Dennett suggests that it would be a waste of resources for the brain to work in this fashion. Once a sensory discrimination has been made, there is no need for it to be made again by other parts of the brain. In other words, the process from the retina to the visual cortex *is* the process by which the colour is discriminated so there is no need for a 'colour box'. The process would have been wasted if it simply supplied its data to a box where the colour to be discriminated was reproduced. But this brings us back to the question: Where is the *blue* that we experience? There is nothing but wavelengths of electromagnetic radiation in the non-mental world, and there is no box filled with blue colour inside our heads so why do we have the blue sensation? Will it ever be possible to use physicalist functionalism to

account for qualia? Leibniz said:

*Supposing that there were a machine whose structure produced thought, sensation, and perception; we could conceive it as increased in size with the same proportions until one was able to enter into its interior as he would into a mill. Now, on going into it he would find only pieces working on one another, but never would he find anything to explain perception (Leibniz, quoted in Rey 1997: p313).*

In this passage, Leibniz is pointing to the mysteriousness of consciousness and the idea that the mere physical and functional organisation of a brain cannot explain the qualities of experience. This leads us to wonder if qualia *supervene* on the physical and functional organisation of the brain. If qualia do not supervene on the physical and functional states of the brain, this would mean that even if we had a complete understanding of the brain's physical composition and functional organisation, we would not be able to account for qualia because they cannot be captured by a physicalist account of the mind. So we could imagine a situation in which the brain's physical and functional structures remain the same and yet qualia do not exist. There are many thought experiments to show the conceivability of an entity functionally the same as us yet being devoid of qualia. Such thought experiments imply that consciousness is a property that cannot be accounted for by physicalism. David Chalmers is a supporter of this view and believes that consciousness cannot be reductively explained. In 'The Conscious Mind', Chalmers attempts to show that no physical explanation of the brain can account for consciousness and for this reason, he holds that consciousness escapes any attempt at a reductive explanation (Chalmers 1996: p93). Chalmers claims that the logical possibility of entities that are functionally the same as us but devoid of consciousness, is one way of showing that consciousness is *not logically supervenient* on the physical composition or functional structure of the brain. The entities that Chalmers refers to are known as zombies.

### **3: ZOMBIES**

A philosopher's zombie is a hypothetical entity that is both *physically* and *functionally* identical to a normal human being. Zombies are molecule for molecule copies of humans and yet they are totally devoid of conscious experiences. My zombie twin is physically identical to me and is therefore functionally identical to me. We can imagine that he lives on a duplicate planet Earth and has led a life that is exactly the same as the life that I have lived. The important point to note is that for Chalmers, my zombie twin is also identical to me with regards his *behaviour* and *psychology*. Psychology, for Chalmers, refers to the functional organisation of the mind. This includes behavioural links and issuing verbal reports. A psychological state is defined by the role it plays in the collective cognitive

system and the behaviour that it gives rise to. For example, if my zombie twin looks at the blue sky, he will make the same comments about his experience as I would. He may state that he enjoys eating chocolate but is hesitant to do so because of the pain he experiences from a hole in one of his teeth. The fact that my zombie twin is *functionally* and *psychologically* identical to me follows from the fact that he is *physically* identical to me. Our brain states play exactly the same functional roles and give rise to an identical psychology. If a certain brain state has the functional role of making me wince in pain, then my zombie twin will behave in exactly the same way when his brain is in that state.

But what about conscious experiences? This is where my zombie twin and I differ. My zombie twin is similar to me in all respects except that his life contains no phenomenal feel. There is nothing it is like to be a zombie. As David Chalmers puts it '*All is dark inside*' (Chalmers 1996: p96). Accepting the possibility such zombies leads to the view that consciousness is not logically supervenient on the physical. This view is closely related to the doctrine of *epiphenomenalism*, which claims that consciousness has no causal powers and has no effect on the world (Guzeldere 1997: p41). If we were to accept the epiphenomenalist view of qualia, we would have to accept that qualia cannot be defined in terms of a role that they play in the formation of beliefs and the production of behaviour. Qualia would seem to be nothing more than a needless feature of human experience, which nature has provided us with for no reason whatsoever. But if qualia serve no purpose, why would nature provide us with them? Before accepting epiphenomenalism, we should explore Chalmers' claims and decide whether or not zombies are possible.

### 3.1: The Logical Possibility of Zombies

For Chalmers' argument, the crucial point is that zombies are logically possible and there is no contradiction or conceptual problem with imagining their existence. Chalmers admits that in the natural world it is likely that anything physically identical to us would have conscious experience just as we do and if a creature was devoid of qualia, it would probably differ from us in some important physical respects. However, even if zombies are *naturally* impossible, their *logical* possibility can still be used to argue for the logical non-supervenience of qualia on the physical. Examining the logical possibility of zombies can also be used to establish physicalist conclusions about qualia. If it could be shown that zombies are not logically possible, we would then be able to pursue reductive functional explanations of consciousness.

It seems that establishing logical possibility is a simple task. For example, we can easily conceive of starships that can traverse the galaxy in just a few hours by imagining a possible world in which humans have developed the technology to make this possible. In the natural world travelling across

the galaxy in such a short space of time would not be possible because to do so we would need to violate Einstein's mass/energy equations. In order to travel across the galaxy in just a few hours, a starship would need to travel faster than light and as such would require an infinite amount of energy. Nevertheless, it is possible to imagine a world in which the laws of physics are different in such a way as to allow such starships to exist. We can use the same technique to establish the logical possibility of many things that do not exist in the natural world. It certainly seems conceivable that there is a possible world in which zombies exist just as it is conceivable that a world exists in which starships can travel faster than light. Chalmers believes that there is nothing incoherent in imagining such a world and he maintains that the burden of proof lies on those who claim that zombies are not logically conceivable (Chalmers 1996: p96). Someone who states that zombies are not logically possible must show that the concept of a creature being both physically and psychologically identical to us but lacking in conscious experience is contradictory. If no such contradiction can be found, we must conclude that consciousness does not logically supervene on the physical.

### *3.2: Assessing the Imaginability of Zombies*

Chalmers often characterises the zombies' lack of conscious experience by using the phrase '*all is dark inside*'. Humans in this world, on the other hand, have a full and rich conscious life and experience sounds, smells, touches and colours. Given the distinction between zombies and humans, would it be correct to say that for humans '*all is colourful inside*'? Probably not. As I stated earlier, according to Dennett, there does not appear to be any place in the brain that contains colour or any other quale. The human brain, like a zombies brain, is dark inside. This, of course, expresses scepticism about the very existence of qualia. Chalmers believes that for humans, there *is* a 'mental' box that contains instances of qualia but in zombies, this box does not exist. When Chalmers speaks of darkness inside a zombie, he is using the term darkness to represent the zombies' total lack of qualia. But this is misleading in some ways and betrays some fundamental problems in imagining the possibility of zombies. Can we really imagine zombies as being all dark inside? Surely the experience of *darkness* or lack of qualia is a quale in itself (Cottrell 1996). Attempting to imagine a zombie involves imagining 'what it would be like' to be a zombie and have no conscious experiences. But thinking of a zombie in these terms cannot work because according to the zombie hypothesis 'there is nothing it is like to be a zombie'. How can we imagine what it is like to be something when there is nothing it is like to be it? There is no conception of darkness or of a lack of experience for a zombie - there is nothing at all going on in there. Imagining what it would be like to be a zombie is not at all like imagining being blind or deaf because presumably there is 'something it is like' to be blind or deaf.

Trying to imagine what it is like to be a zombie is like trying to imagine what it is like to not exist, and I maintain that this is an impossible task. The obvious reply to this would be to ask why it is necessary to imagine what it is like to be a zombie. People can imagine unicorns as being logically possible without having to imagine what it is like to be a unicorn, so why can't the same be true of zombies? I would respond to this by suggesting that there is a difference between imagining a unicorn and imagining a zombie. To imagine a unicorn, one simply has to take two naturally existing things in the world and join them together - a horse and a horn. We can imagine a world in which evolution gave horses horns because we see creatures with horns in our world all the time. But imagining our zombie twins is more difficult. We can imagine our twins as being identical to us easily enough, but we also have to imagine that they have no conscious experiences. To accomplish this, we have to imagine ourselves as having no conscious experiences and this is where the problem lies. We cannot imagine what it is like to have absolutely no conscious experiences. This is why philosophers such as Daniel Dennett believe that people inevitably fail at the task of imagining zombies. Often the assumptions people make about the zombies' lack of an internal life is incompatible with the statements they make about the zombies' external behaviour. Their description of zombies often contradicts the idea that zombies are behaviourally and functionally identical to us (Dennett 1995a: p172). Indeed, as I have already explained, it seems to be impossible to imagine completely normal external behaviour at the same time as imagining a total internal absence of experience. So it would seem that one of the major problems with imagining zombies is that people underestimate exactly what is involved and they often include facts that are not really imaginable. Many of the properties ascribed to zombies are *stipulated* rather than *imagined*. The problem here is that it is possible to make *any* stipulation or assertion that we feel is necessary to make a thought experiment work even though these stipulations may be totally unimaginable (Horgan 1987: pp496-500). For example: Imagine a creature that is totally identical to yourself in every physical and psychological respect. Imagine also that this creature has a full conscious experience that is identical to your own conscious experience. Now imagine that this creature is not alive (Dennett 1995a: p176). Given that this creature is physically, psychologically and behaviourally identical to us, we may think that we have discovered a mysterious property that cannot be explained. Life! We are alive but our imagined twins are not alive so what is this thing called life? Dennett suggests that this mystery would only arise if we thought that life was some extra property that could be removed or added to the totality of a normally functional human (Dennett 1995a: p176). But regardless of what we think, life is not this sort of property. Perhaps we should not think of consciousness as some property that could be removed from the totality of human function. Some people may claim to be able to imagine such a creature as described above, but I suspect that their imaginations would fail to fully conceive of such a scenario. I believe that if people claim to have imagined duplicates of

themselves that are not alive, they have not actually imagined the situation, they have only made stipulations about what the creature should be like. Perhaps the same is true when we attempt to imagine zombies.

Shoemaker has pointed to another problem with the conceivability of zombies and has used the zombie type argument to argue for reductive explanations of consciousness. As humans, we make judgements and form beliefs about our experiences and qualia. We can do this through introspection. If this is true, it would seem that our experiences *cause* our introspective beliefs. Now, remember that zombies are physically, functionally, psychologically, and behaviourally identical to humans. This means that zombies and humans share identical introspective mechanisms. Of course, since zombies exhibit the same behaviour as humans, it follows that they make the same claims about their experiences as we do. They talk about enjoying a rich and full conscious life in exactly the same way as humans do. Zombies form beliefs about their experiences through their introspective mechanisms, and since they have no qualia, we must accept that qualia do not play a role in the formation of their beliefs. But, if we have the same introspective mechanisms, then it would seem that qualia do not play any causal role in our formation of beliefs. Zombies insist that they are not zombies, just as we do. But if we are forming those beliefs by using exactly the same mechanisms as zombies, then we should not be able to tell whether or not we are zombies (Shoemaker 1975: pp291-315). Qualia, for humans, seems to be accessible to introspection and therefore seems to play a role in our formation of beliefs and subsequent behaviour. If zombies have the same mechanisms as we do, then given that they have no qualia, they should behave in a different way. But according to the argument, zombies behave in exactly the same way as humans. This is a contradiction, and as such, highlights a very important problem with the logical possibility of zombies. Chalmers suggests that a response to this problem would be to argue that causal theories of knowledge and belief acquisition are not appropriate for accounting for our knowledge of consciousness (Chalmers 1996: p193). By taking a dualist stance, we could claim that knowledge of conscious experience is fundamentally different to other forms of knowledge, and that our knowledge of conscious experience is not brought about by any sort of causal relationship between knowledge and experience. Our knowledge of conscious experience consists in another sort of relationship. But Chalmers does not offer this alternative.

*"... there is good reason to believe that the epistemology and semantics of experience cannot be essentially causal, and should instead be understood in other terms. ... A full understanding of these issues would require a lengthy separate investigation ..."* (Chalmers 1996: p209)

So the mystery remains. Chalmers has replied to Shoemaker's objection by claiming that causal theories cannot account for our knowledge of experience, but he offers us no alternative theory. He merely suggests that we might go about accounting for knowledge by appealing to a

dualist point of view. At this stage I see no reason to reject the causal accounts of knowledge acquisition. Shoemaker's objection should therefore be taken very seriously.

Despite the problems associated with the conceivability of zombies, philosophers such as Chalmers maintain that the concept of zombies is perfectly coherent and leads to no logical contradiction. But *why* do people think it is so easy to imagine zombies? Allin Cottrell suggests that the attraction in thinking of zombies as conceivable lies in our view of human cognition as an information processing system (Cottrell 1996). Many contemporary theories of mind make the claim that the brain is an information processing system and is in essence some sort of biological computer. These theories assume that our mental states should be understood as functional states, which are to be understood as connectionist or symbolic computational states. If the brain is to be thought of as a computer, it is very easy to imagine the brain carrying out its computational tasks without any qualia or subjective experience, just as it is natural for us to assume that computers have no internal experiences. In fact, it would be difficult to conceive of a present day computer as having subjective experiences and enjoying the raw feels we experience.

Computational theories of mind may offer some insight into why some believe it is easy to imagine zombies, but I think that imagining zombies in this fashion falls short of capturing the complexity of the human brain. It is probably true that the brain is nothing more than a sophisticated information processing device, but appealing to the computational nature of human cognition should not make it any easier to imagine zombies. We assume that computers have no qualia, but zombies are more sophisticated than electronic computers. Zombie brains are identical to human brains in every respect. Human brains do not work in the same way as the other computers we come into contact with. Imagining desktop PC's as being devoid of qualia does not help us imagine creatures with human brains who lack qualia because they are totally different types of systems. On the otherhand, imagine that a computer is constructed that works in exactly the same fashion as a human brain. If this system was designed as a copy of an existing biological brain and was installed in an android body, it would be hard to imagine that it did not have qualia because like a zombie, it would have the same psychology and behaviour as us. If it could be shown that such a device was devoid of qualia, we would then have to discover how to add qualia to the system that was already complete. This would lead us to an epiphenomenal position, which implies that it would be possible to remove the qualia from a human subject while keeping all of her physical and psychological functions in place.

Given the difficulties in adequately imagining the possibility of zombies it would seem that zombie supporters cannot be justified in stating that zombies are logically possible. If, however, a zombie supporter could find

an existing case of zombiehood, the case for zombies would become convincing. As we have seen there would be difficulties in classifying an entity as a zombie because we only have access to its verbal claims about consciousness and these claims would lead us to believe that it was conscious in the full human sense. But all is not lost for the zombie supporter yet. Examples have been found of people who could be considered to be *partial* zombies. These people are missing the subjective conscious experience associated with a particular aspect of cognition and can verbally report upon its absence. In the literature, the most cited example of this type of deficit involves the curious phenomenon known as 'blindsight'.

#### **4: Blindsight! Visual Zombies on Earth**

Lesions to the visual cortex can cause a neurological condition known as a scotoma (a loss of vision in the corresponding area of the visual field). People who have this condition retain most of their normal vision, but are aware of a blind spot in their visual field. The interesting feature of this disorder is that in some cases, there is evidence that some visual processing is going on within the blind area of the visual field. Experiments have been carried out in which patients have had a stimulus (a flash of light) presented to their blind region and were asked if they had seen anything. Obviously, because they are blind in that area, the patients reported that they were unaware of the flash of light. However, when forced to take a guess and point to where the flash occurred, many subjects exhibit surprising accuracy. Somehow these people are detecting stimuli that are presented to their visual field. These experiments have been carried out many times and there is now little doubt that these subjects are performing much better than chance. It seems that in these cases, the visual system is carrying out some of its normal information processing tasks without the patient's conscious awareness. There is vision without qualia. This phenomenon is known as blindsight and has been used, by many, as a *natural* example of partial zombiehood (Block 1995: p385). The fact that blindsight can be used to show that it is possible for people to see things without the phenomenal feel associated with sight makes the phenomenon very attractive to many zombie supporters. Blindsight cases certainly seem to show that (in the case of vision at least) consciousness is not an essential part of the cognitive process that goes on within the brain. However, despite its intuitive appeal, the phenomenon of blindsight has not convinced everyone that partial zombiehood is possible. The first problem with using blindsight as an example of partial zombiehood should be obvious. One of the zombie argument's main premises states that zombies should be *physically and functionally identical* to human beings. But blindsighted patients have suffered physical damage to part of their cortex and their brains are no longer physically identical to their 'predamaged' brains. Using the

impaired visual processing systems of a blindsight patient to count as an example of partial zombiehood constitutes a failure in recognising the implications of such an exercise. One may think that one has found an example of partial zombiehood, but this would be failing to take into account the difference between a blindsighted person and a fully functional human. This in itself does not refute Chalmers' claim that experience is not logically supervenient on the physical. It does, however, show that experience is naturally supervenient on cortical activity.

Gerald Vision (1998: p137) claims that arguments that use blindsight as evidence of vision without phenomenology usually rely on misconceptions about blindsight. Daniel Dennett (1993: p331) makes similar claims. He suggests that philosophers who are reliant on blindsight as an example of visual perception without consciousness are overlooking certain basic facts about the phenomena. In some cases, it could be possible that the blindsight subjects are receiving unintended hints from the experimenter and this could explain the fact that sometimes subjects do better in the experiments than other times. Of course, in an experimental setting, these things are carefully controlled, but we can never fully dismiss the possibility of unknown cues from the experimenter.

Dennett has suggested that the blindsight phenomenon does not show that blindsighted people lack visual consciousness. For Dennett, blindsighted people have an impoverished visual content within their blind region, and as such, information contained in that region has a very low degree of influence on behaviour (Dennett 1995b: p417). This would explain why blindsighted patients do not spontaneously offer reports about the content of their blind spot. The visual content is not strong enough to have that type of influence on their behaviour. They have to be cued from an external source. Dennett claims, however, that it would be possible for a blindsighted patient to be trained to guess correctly *when* to make a guess about the contents of her blindspot. If she still performed above chance in the experiments, she would *qualify* as being visually conscious. Such a person could treat her visual stimuli in the same way as we treat conscious experience. She could think about and report on the stimuli within her visual field and for all intents and purposes be conscious in the sense that consciousness consists in reportability or an availability for deliberation. Presumably in this scenario, the content of the blind region would have a strong enough influence on behaviour to allow the blindsighted patient to respond to cues from other centres in the brain. But would this count as visual consciousness? What about the phenomenal feel of visual experience? Gerald Vision has suggested that blindsighted patients would not be able to become visually conscious as Dennett suggests. Blindsight patients have lesions that block crucial inputs to the *ventral* system, which is involved in object recognition. It is unlikely that the neural pathways supporting visual recognition could be regrown by simply training a patient to recognise when to guess (Vision 1998: p153). This point is obvious. Training a patient to guess when to guess about the content of their blind region would not be sufficient to repair damage to

neural pathways. If a patient *was* successfully trained to recognise when to guess, visual consciousness would still be missing. Such training would probably involve recognising non-visual cues such as muscle readiness.

Vision claims that arguments from blindsight usually involve misconceptions about the phenomenon. It is often assumed that blindsight patients have some sort of visual beliefs about the features of the objects within their blindspot and that these beliefs can be discovered through verbal reports of blindsight patients. This assumption is misguided, however, because evidence has shown that the success rate of verbal responses to stimuli presented to the blind region is only slightly above chance. In experiments carried out by Zihl and Von Cramon, blindsight subjects had a light flashed in their blind regions and were asked to respond in three different ways: by pressing a button, blinking or saying "yes". While the success rate for pressing the button or blinking was well above chance, the subjects performed significantly worse with the verbal response (Vision 1998: p146). Blindsight patients can reach and grasp at objects and can even rotate their hand in relation to the orientation of an object, however they usually cannot issue any verbal report about the object. In a simplified account of the visual process, Gerald Vision draws a distinction between two functionally discrete streams of visual processing - the ventral stream and the dorsal stream. The ventral stream is responsible for computing object features such as colour, shape and motion, while the dorsal stream processes information with regards to location. The dorsal stream is used for immediate responses to stimuli such as grasping for food or avoiding obstacles while the ventral stream is responsible for object recognition and classification. This distinction may provide an explanation of why blindsight patients retain visuomotor skills, such as reaching, while higher level aspects of vision that may be responsible for verbal reportability are lacking. Vision suggests that in cases where blindsight patients manage to achieve higher than chance success in verbal response, patients may be relying on clues from premotor readiness in their muscles (Vision 1998: p151). This would imply that the verbal reports of blindsight patients relate only to muscle readiness rather than what is being visually perceived. For Vision, blindsight is not the result of dissociations within a single unified visual system, its occurrence can be explained more adequately by considering the interactions between distinct self contained systems.

An understanding of blindsight shows us that the phenomenon cannot easily be used as an example of partial zombiehood. Zombie arguments are supposed to convince us that normal behaviour is possible without qualia and that consciousness is not logically supervenient on the physical. Blindsight, on the otherhand, shows us that physical changes to the brain can affect certain aspects of vision and that in humans, the visual process probably consists of interactions between several separate systems, some of which continue to function normally after damage has occurred to the brain. The point is that blindsight, while an interesting phenomenon, does not provide an example of physical and behavioural sameness without

qualia. It is true that in blindsight, the visual qualia are missing, but this is the result of physical and psychological changes within the brain and so the blindsight phenomenon cannot be used to argue for the logical non-supervenience of qualia on the physical. If anything, blindsight shows that the physical state of the brain *is* an essential condition for visual qualia. This, however, is not sufficient to show that the qualia themselves are physical, it merely shows us that there is a causal link between physical processes and mental states.

#### *4.1: The Absent Minded Truck Driver - A Part-time Zombie*

There is an example, which we may all be familiar with, that can serve as an instance in which consciousness seems to be missing for short periods of normal waking time. Imagine that you are a long distance truck driver. After driving for long periods of time, it is quite common to suddenly become aware of the fact that for some time you have been driving without being aware of what you have been doing (Armstrong 1981: p723). You may have been on a long trip and suddenly realise that you have no recollection of events that have taken place during the last part of the trip. It is a sobering experience that most drivers have had. It even happens on short trips. You find yourself arriving home without any memory of your actions during the trip. But during the trip, you were engaged in high level control operations including motor control and decision making. You were negotiating traffic signals, avoiding cars, and changing gears, but lacked consciousness of these actions and events. Would it be correct to say that this is an example of part-time zombiehood? David Armstrong believes that the driver in this scenario had minimal consciousness. This is to say that the truck driver was perceptually conscious of the road, and that there was mental activity involved in guiding the driver's actions. For Armstrong, what the driver lacked was introspective awareness - an awareness of the activities going on in her mind during the trip. In this sense, the driver's actions were driven by a primitive level of mental functioning. This mental functioning could be considered to be 'unconscious' (Armstrong 1981: p724). Fred Dretske disagrees with this claim and believes that the only sense in which the driver's mental states were unconscious, is that the driver was not conscious of having those states. This does not mean that the states themselves were unconscious (Dretske 1993: p778). Dretske makes a distinction between consciousness of *objects* and consciousness of *facts*. For Dretske, the driver could have had conscious experiences of objects - for example green traffic lights - without having any consciousness of the facts about those objects. We can have a conscious experience *of* a green light without having a conscious belief *that* it is a green light. I think Dretske is correct and this shows that the case of the absent minded truck driver cannot be used as an example of zombiehood.

Another way of explaining the case of the absent minded truck driver

involves memory. It could be the case that the experiences of the truck driver were not encoded in long term memory, and that is why she cannot introspect and make claims about the long journey. It may seem to the truck driver that 'there was nothing it was like' to be on the long drive, but that thought occurs after the event. If we were to probe the driver at any point during the trip, we would probably get reports of a rich and full conscious experience. The point is that if the truck driver cannot remember conscious experiences, it does not follow that they did not exist. I cannot remember any of my conscious experiences from Wednesday the 1st of October 1997, but that does not mean that I was a zombie on that day. Zombies are defined as being identical to humans in every respect, except that they lack qualia. In the case of the absent minded truck driver it is far from clear that the driver lacked qualia. I think the alternative ways of describing what happens in this case are more convincing than the claims that the driver had no conscious experience of her trip.

### **What has Become of the Zombies?**

The possibility of zombies has been asserted by philosophers in an attempt to show that consciousness is not logically supervenient on the physical structure of the brain. Zombies are physically, functionally, behaviourally and psychologically identical to humans and yet they have no qualia. In order to construct arguments that appeal to the possibility of zombies, the zombie supporter must first show that zombies are logically possible. For David Chalmers, that simply involves imagining a possible world in which zombies exist. He claims that the burden of proof lies with those who wish to deny the logical possibility of zombies. But, we have seen that imagining zombies is not as easy as Chalmers has lead us to believe. Our imaginations inevitably fail at the task. We cannot imagine a zombie. We can only make stipulations about what features a zombie might have (or lack). Furthermore, Shoemaker's argument has shown us that the definition of a zombie leads us into a contradiction. Qualia, for humans, are accessible to introspection and therefore seem to play a role in our behaviour. If zombies have the same introspective mechanisms as we do, then given that they have no qualia, they should behave in a different way. But according to the argument, zombies behave in exactly the same way as humans.

An option for the zombie supporter might have been to find a case of zombiehood in the real world. If this could be done, the zombie arguments would still work. The problem is that there does not seem to be a clear cut case of zombiehood in our world. For both the case of blindsight and the absent minded truck driver, we could easily find alternative explanations for the phenomena without resorting to the conclusion that they provide us with true examples of zombiehood.

With no existing examples of zombiehood, the zombie supporter must appeal solely to conceivability arguments. But after highlighting the problems with imagining zombies, it would seem that the zombie supporter has a lot of work to do.

I cannot show that consciousness *is* logically supervenient on the physical, though I think it is. I have, however, provided reasons for which arguments involving zombies cannot show that consciousness is *not* logically supervenient on the physical. The zombie arguments remain unconvincing and do not offer any serious challenge to physicalist descriptions of the mind.

## References

Armstrong, D. (1981), "What is Consciousness" in *The Nature of Consciousness*, Edited by Block. N., Flanagan. O., and Guzeldere. G., MIT Press, A Bradford Book, 1997.

Block, N. (1995), "On a Confusion about a Function of Consciousness" in *The Nature of Consciousness*, Edited by Block. N., Flanagan. O., and Guzeldere. G., MIT Press, A Bradford Book, 1997.

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

Cottrell, A. (1996), "On the conceivability of zombies: Chalmers v. Dennett", <http://ling.ucsc.edu/~chalmers/zombies.html>, 1996, (Downloaded 1/7/98).

Dennett, D. (1988), "Quining Qualia" in *The Nature of Consciousness*, Edited by Block. N., Flanagan. O., and Guzeldere. G., MIT Press, A Bradford Book, 1997.

Dennett, D. (1993), *Consciousness Explained*, Penguin Books, 1993.

Dennett, D. (1995a), "The Unimagined Preposterousness of Zombies" in *Brainchildren*, Penguin Books, 1998.

Dennett, D. (1995b), "The Path Not Taken" in *The Nature of Consciousness*, Edited by Block. N., Flanagan. O., and Guzeldere. G., MIT Press, A Bradford Book, 1997.

Dretske, F. (1993), "Conscious Experience" in *The Nature of Consciousness*, Edited by Block. N., Flanagan. O., and Guzeldere. G., MIT Press, A Bradford Book, 1997.

Guzeldere, G. (1997), "Introduction: The Many Faces of Consciousness: A Field Guide" in *The Nature of Consciousness*, Edited by Block. N.,

Flanagan. O., and Guzeldere. G., MIT Press, A Bradford Book, 1997.

Horgan, T. (1987), "Supervenient Qualia" in *Philosophical Review* Vol xcvi, no.4, October 1987, Edited by Wood, A., Irwin, T., Appiah, A., Cornell University.

Levine, J. (1993), "On Leaving Out What It's Like" in *The Nature of Consciousness*, Edited by Block. N., Flanagan. O., and Guzeldere. G., MIT Press, A Bradford Book, 1997.

Rey, G. (1997), *Contemporary Philosophy of Mind*, Blackwell Publishers, 1997.

Shoemaker, S. (1975), "Functionalism and Qualia" in *Philosophical Studies* 27, 1975.

Vision, G. (1998), "Blindsight and Philosophy" in *Philosophical Psychology* Vol 2, Number 2, June 1998.