

The Inconceivability of a Zombie World

Bharath Vallabha

I. Introduction

Imagine that there is a world physically identical to the actual world. All the atoms, stones, bodies and other physical entities are exactly the same in the two worlds. Imagine further that the two worlds are identical for all time. The occurrence of the big bang, the formation of our solar system, the separation of the continents on earth and so on are exactly the same. Now can you imagine that there is nonetheless a difference in conscious states in the two worlds? Let us call *a zombie world* one which is physically identical to the actual world but in which no one is conscious. Is a zombie world conceivable?

David Chalmers says yes.¹ He argues that conscious experiences are had from a first-person perspective which is characterized by what it is like for the subject to undergo the experience. When you see the vibrant colors of a sunset or eat a piece of chocolate cake, there is something it is like for you to have those experiences. He claims it is conceivable that though your twin in the zombie world behaves exactly like you do, there is nothing it is like for him or her to see a sunset or have any other experience.

Chalmers makes this claim as part of an argument for dualism. According to Chalmers, dualism is the denial of physicalism, which is the view that any difference in a state of the world implies a difference in basic physical states. In addition to the conceivability of a zombie world, Chalmers argues that conceivability implies possibility. So it is possible that there can be a difference in conscious states without a difference in basic physical states and therefore dualism is true.

Most physicalists have responded to Chalmers by conceding that a zombie world is conceivable but denying that conceivability implies possibility.² They argue that, for example, we can conceive that water might not have been H₂O, but that it is nonetheless a necessary, a posteriori truth. Though this response to the zombie argument raises

¹ See Chalmers 1996, pgs. 93-122.

² See Block and Stalnaker 1999, Hill 1997, Levine 1983 and Loar 1997.

interesting metaphysical issues, I think it concedes the most important point to dualism. Denying the possibility of a zombie world defends the letter of physicalism, but accepting its conceivability leaves untouched the intuitive appeal and the challenge of dualism.

Some physicalists deny the conceivability of a zombie world by pursuing the representationalist strategy.³ They argue that conscious experiences are representational and that representational states are functionally defined in terms of their causal relations to other representational states and the environment. They also argue that the zombie world is a causal duplicate of the actual world and so it contains conscious experiences. I think this response to Chalmers has led to a standoff. Both Chalmers and the representationalist agree that there is a sense of consciousness in which it can be functionally defined. But does consciousness in this sense exhaust what it is to be conscious? Here the intuitions of Chalmers and the representationalist simply differ. For Chalmers it is an obvious truth of experience that there is consciousness in a further, first-personal sense. The representationalist denies any such further sense.⁴

In this paper I will pursue a different strategy for denying the conceivability of a zombie world. Instead of focusing on the relation between representational states and consciousness, I will focus on the relation between *actions* and *consciousness*. My argument has two premises. First, a world physically identical to the actual world contains agents: it contains organisms which perform intentional physical movements. Second, agents are conscious beings: performing intentional physical movements requires consciousness.

My aim is not to reduce consciousness to actions and actions to causal relations in the physical world. So actions do not play the same role in my argument that representational states play in the representationalist's argument. My aim rather is to show that two things are essentially true of actions: they are a part of the physical world and they presuppose consciousness. This is why a zombie world is inconceivable. This doesn't imply that consciousness isn't currently a mystery to us. I think it is mystery, for even the best current theories of it seem only to scratch the surface. But it does imply that

³ See Dennett 1991, Dretske 1995 and Harman 1990.

⁴ See Chalmers 2002, pgs. 251-253. The dialectic is complicated by the fact that Chalmers believes that the sense of consciousness which can't be functionally defined is nonetheless representational. See Chalmers 2004 and 2005.

consciousness is a mystery in part because we are also in the dark about the nature of actions. In particular, it is a mystery how actions are both a part of the physical world and presuppose consciousness.

The structure of the paper is as follows. In section 2 I will present Chalmers' view. In section 3 I will give an overview of my argument. In section 4 I will argue that there are agents in the zombie world and in section 5 that agents are conscious. And I conclude in section 6.

II. Chalmers' View

Chalmers claims that it is conceivable that there is a world physically identical to the actual world but which lacks consciousness. What does he mean by the physical world and consciousness?

By the physical world Chalmers means a world with basic physical facts and all the facts they imply. By basic physical facts Chalmers means the facts of a completed physics. They include two things: facts about basic physical entities such as quarks and strings or their analogues in a completed physics; and facts about the basic physical laws. Non-basic physical facts are facts implied by the basic physical facts. So facts about ordinary physical objects such as rocks, trees and chairs are non-basic physical facts because they are implied by basic physical facts.

A world physically identical to the actual world is one which has all the basic physical facts of the actual world and so all the facts which are implied by those basic facts. For example, suppose in the actual world there is a mountain at a certain place and time. Then in the duplicate physical world, there will also be a mountain with the same properties at that place and time. For given all the facts about the basic physical entities in the actual world and the facts about the basic physical laws, it is not possible that there can be a difference in non-basic physical facts concerning the mountain. There is no possibility for a gap, as it were, between the facts about the basic physical entities and laws and facts about the mountain.

The same point applies to all non-basic physical entities, including the bodies of people and social artifacts such as buildings and art works. If in the actual world a person is standing in front of the *Starry Night* by Van Gogh at a certain place and time, then in

the duplicate physical world there will be a person with identical physical properties standing in front of a physically identical painting at that place and time. The most complicated physical details of that person's body will be duplicated, including the states of his brain and the causal history of his birth, childhood and so on. For just as with the mountain, if all the basic physical facts of the actual world are duplicated, even the most complicated physical facts of the person and his environment are going to be duplicated.

But does this mean that the person's consciousness is also going to be duplicated? Or is it conceivable that though the person in the actual world is conscious of the *Starry Night* and its surroundings, the duplicate person is not conscious at all?

To explain the sense of consciousness at issue here, Chalmers distinguishes between *psychological* and *phenomenal* consciousness. Psychological consciousness is characterized by its causal relations to inputs from the environment, other mental states and behavior. So it is defined functionally in terms of its causal role in the cognitive economy of a subject. Phenomenal consciousness in contrast is characterized by what it is like to have the conscious experience. It is defined in terms of the qualitative properties of how it feels for the subject to be conscious.

When the person is seeing the *Starry Night* he is psychologically conscious of it if he is gaining information about it. Through that information he forms beliefs, remembers other things it is related to, learns about the painting and in general is able to intelligently act with respect to it. He is thus conscious of the painting in that he is responsive to it in certain ways. He is phenomenally conscious of the painting if there is something it is like for him to be seeing the painting. He then experiences the vivid blue which dominates the painting and feels the tremors of elation or terror which it evokes in him. He is conscious of it not just in terms of gaining information but by having experiences which feel a certain way.

According to Chalmers the person in the duplicate physical world will be psychologically conscious. Since consciousness in this sense is a matter of how he causally interacts with the world, it is determined by the basic physical facts. If the duplicate world is identical to the actual world in terms of the basic physical entities and the basic physical laws, then it is not possible for the causal properties of the person in the duplicate world to differ from the causal properties of the person in the actual world.

But Chalmers claims that it is conceivable that the person in the duplicate world is not phenomenally conscious. Whereas the person in the actual world experiences the colors of the painting and his resulting emotions, the person in the duplicate world might not experience anything at all. Though externally he is identical to the person in the actual world, it is conceivable that internally he is not conscious at all. Chalmers writes, “[My zombie twin] will be *psychologically* identical to me... He will even be ‘conscious’ in the functional senses... It is just that none of this functioning will be accompanied by any real conscious experience... There is nothing it is like to be a zombie.”⁵

On Chalmers’ view therefore phenomenal consciousness is unique among all the properties of a subject. Whereas all the other properties are determined by the basic physical facts of the world, phenomenal consciousness is not.

III. Outline of My Argument

I think a duplicate physical world without phenomenal consciousness is inconceivable. My argument for this claim will depend on what is required for a physical object to be *an agent*. By an agent I mean an organism which performs actions.

By organisms I mean animals. They are living beings which can move around in the physical world through the functioning of their body in order to obtain food and reproduce. In this sense organisms differ both from inanimate objects such as stones and tables and living beings such as plants which are incapable of locomotion.

By an action I mean an intentional, physical movement. In this paper I am not going to talk about actions such as making a decision about what to eat, imagining the Eiffel Tower or playing a piece of music in your head. We perform these events as intentionally as raising one’s hand or jumping over a puddle. But in this paper I am going to focus only on the latter and by an action I will mean a certain kind of physical movement.

In the next section I will state what makes a physical movement intentional. Here I want to make one point about what I don’t mean by intentional. I don’t mean the feeling of intentionally performing a physical movement. One might think that physical movements themselves are never intentional; they are only the causal effects of an

⁵ Chalmers 1996, pg. 95.

internal experience of willing which is what is really intentional. I am going to assume that this view is false, for to argue against it here would take me too far a field. As I understand it, physical movements themselves can be intentional.

My argument that a zombie world is inconceivable is as follows:

1. *The Embodiment Claim.* A world physically identical to the actual world contains agents.
2. *The Consciousness Claim.* Agents are phenomenally conscious.
3. So, a world physically identical to the actual world contains phenomenal consciousness.

The zombie world seems conceivable when one fails to recognize that agents are essentially related to both the physical world and phenomenal consciousness. In the next two sections I will argue for these relations.

IV. The Embodiment Claim

4.1. *The Argument*

I want to first set aside two tempting arguments for the embodiment claim which don't work. The first argument is that physical movements supervene on basic physical facts and therefore there are actions in a duplicate physical world. For example, when a person is walking around, his legs traverse space, and this would not be possible without a change in the physical entities which constitute the legs. This argument establishes the trivial truth that there are physical movements in the duplicate world. But it doesn't establish that those physical movements are intentional.

The second argument is as follows. Mental states in the psychological sense are defined by their causal relations to inputs from the environment, other such mental states and actions. Since such states are defined causally, they are in the duplicate world and therefore so are actions. But what is here meant by actions? If they are physical movements, it faces the same objection as the above argument. And if they are intentional, physical movements, the argument is circular. For the claim that there are psychological mental states in the duplicate world would then presuppose, instead of establishing, that there are actions in that world.

Here is my argument for the embodiment claim:

1. A world physically identical to the actual world contains organisms.
2. *The embodied theory of action.* The physical movement of an organism is intentional if and only if (a) it is caused by a perception of an object in the environment, (b) the physical movement is part of an extended, reciprocal causal relation between the organism and the object, and (c) the parts of the organism relevant to the causal relations in (a) and (b) are healthy.
3. If a world is physically identical to the actual world, the causes of the physical movements in the two worlds are identical.
4. So, a world physically identical to the actual world contains agents.

(1) states that there are organisms in the duplicate world. (2) and (3) together imply that if there are organisms in the duplicate world, their physical movements are intentional. (2) is the crux of this argument and is most in need of defense. Before defending it, I will briefly state the motivation for the other two premises.

An organism is an animate, locomotive being. The theory of evolution and the discovery of DNA as the mechanism by which the instructions for life are passed on show that an animate being is an enormously complex physical structure. A life form such as bacteria or a plant is composed of basic physical entities organized in a certain way. Moreover, the difference between a non-locomotive and a locomotive animate being consists only of a certain kind of complexity: the latter interacts with more of the physical surface of its environment than the former. So given that a physical duplicate of the actual world contains the same physical complexity as the actual world, it contains organisms.

A world physically identical to the actual world contains the basic physical facts of the latter and the facts which logically supervene on them. Some of the supervening facts concern physical entities which exist: rocks, trees, planets and so on. Other supervening facts concern causal relations between those physical entities: the rolling rock was stopped by the tree, the orbit of the planet was altered by the gravitational force of the sun and so on. Thus in a duplicate physical world the facts concerning the causal relations between physical entities are also duplicated. If a physical movement is caused

in a certain way in the actual world, it is caused in the same way in the duplicate world.

One might object to (1) by saying that the advances of biology don't show that life is a complex physical structure; they show rather that a lot of what we ordinarily take to be life, such as the lower animals, isn't really life. For the mark of life is a free will which determines but isn't itself determined by the physical world. Similarly, against (3) one might say that some physical movements in the actual world have non-physical causes and therefore the physical movements in the duplicate world might causally differ from those in the actual world.

I am not going to address these objections in this paper, since they require a separate discussion of free will. Barring extreme views like these, I think (1) and (3) are intuitive.

4.2. The Embodied Theory of Action

Suppose a lion is resting on the Serengeti plains. In the distance it sees a gazelle approaching and it goes into its stalking position. It gets up, stays low, moves its legs slowly and is all the while intensely focused on the gazelle. As the gazelle comes closer, the lion breaks into a stride and runs after it. In stalking the gazelle, the lion moves its legs intentionally. The legs don't move as knee-jerk reactions or as something out of the lion's control: part of what the lion is doing is moving its legs. What makes these physical movements intentional?

There is an intuitive answer to this question. It is that the lion sees the gazelle, moving the legs is a part of stalking and given that they are hungry or playful, stalking is what lions do when they see animals they can eat. The lion's leg movements are intentional because, as we might say truisitically, the lion is being a lion. I think this intuitive answer is basically correct and what I call *the embodied theory of action* is a crystallization of it. The theory describes the physical structure of action.

The first condition of the embodied theory is that an intentional physical movement is caused by a perception of an object in the environment. An action begins with perception. By perception I mean the intake of information from the world through a sense organ which enables the functioning of the organism. In section V I will connect perception with phenomenal consciousness, but in this section I mean perception in an

informational sense.

Suppose the lion is performing all the physical movements of stalking, but it does not have any perceptions (visual, auditory and so on). The lion does not obtain any sensory information from the world. It looks like it has functioning eyes, ears and so on, but actually they do not process any information. The lion just happens to be organized so as to make stalking like movements. I think in this case the leg movements aren't intentional. They are like the movements of a mechanical horse children can ride to simulate riding a real horse. The legs go up and down and there is a mechanism which causes this, but they are not being intentionally performed.

One difference between the imperceptive lion and the mechanical horse is that the former has a functioning body. So couldn't its physical movements be intentional just in virtue of being caused by a functioning heart, lungs and so on? No, for without sensory information from the world, it will just be a fluke that there are stalking like movements. An intentional physical movement doesn't just transform the environment: it transforms it so that the organism can gain from the environment something it is otherwise unable to obtain. An intentional physical movement is therefore the organism's reaction to the environment. But this is not possible if the organism has no sensory information. The lion and the horse are oblivious to the environment and so though they can cause it, they do not act upon it.

The second condition of the embodied theory is that the physical movement is part of an extended, reciprocal causal relation between the organism and the perceived object. The first condition is a constraint on how the physical movement is caused. The second condition is a constraint on how the movement is part of an extended causal structure.

Suppose the lion sees the gazelle, moves one leg in its direction and then abruptly stops because it has lost interest in the gazelle. Imagine further that whenever the lion sees a gazelle, it always performs this aborted movement. Seeing the gazelle makes the lion move in its direction, but just as soon as it performs the movement it turns and walks away from the gazelle. I think in this case the leg movement isn't intentional. The lion moves its leg after seeing the gazelle in just the way a person lifts the leg after the doctor taps the knee. It's a response which simply happens.

It is therefore not enough for a physical movement to be intentional that it is caused by a perception of an object. In fact, a stronger claim is true: no matter what the intermediaries in the causal chain between the perception and the movement, this causal chain is not enough to make the movement intentional. For it is essential for a intentional physical movement that the question “why?” in a certain sense is applicable to it.⁶ The answer to this question gives the purpose of the movement. For example, I am walking in order to go to the store or I am raising my hand in order to vote. If a physical movement has no purpose in this sense, it is not intentional. The question “why?” might then still apply in a different sense (such as why is the hand going up at this speed?) but the question will apply only in this sense.

The fact that an intentional physical movement has a purpose means that the movement is directed into the future. Thus the causal history of the movement cannot be sufficient to explain what makes a physical movement intentional. The causal history by itself answers the “why?” question only in the non-purposive sense and leaves unaddressed the “why?” question in the purposive sense. This is true even if the causal history involves cognitive states such as perceptions, beliefs and desires. If the lion’s desire to eat the gazelle is part of the causal history of its leg movement, there is still the issue of how the desire relates to the purpose of the movement. Perhaps fulfilling the desire is the purpose of the movement. But then the purpose consists of eating the gazelle and not of the fact that the desire was a cause of the movement.

In what way must a physical movement be directed into the future in order to be intentional? Suppose the lion sees a gazelle, moves one leg in its direction and then performs stalking movements seemingly directed at another gazelle. I think in this case the preliminary physical movement isn’t intentional because it has no purpose. If it has a purpose, it must consist in interacting with either the first or the second gazelle. But with regard to the first gazelle this case is like the previous one where the lion lost interest in the gazelle; the fact that the movement is caused by a perception of the gazelle doesn’t make it intentional. With regard to the second gazelle, this case is like the earlier case of the imperceptive lion; since the physical movement isn’t caused by a perception of the second gazelle, it isn’t intentionally directed towards it.

⁶ See Anscombe 2000, pgs. 9-11.

Therefore a physical movement is intentional only if it is directed towards the same object the perception of which caused the movement. This doesn't mean simply that the movement causes a change in the perceived object. In a minimal sense that is true of all the objects in the organism's environment and so it doesn't specify the unique relation between the organism and the perceived object.

It means rather that a physical movement is intentional if it is a part of the following extended causal interaction between the organism and the object. The object causes a perception of the object, which together with the organizational state of the organism causes a physical movement, which in turn causes the organism to better perceive the object (that is, sensorily gain more information about it), which again together with the organizational state of the organism causes a physical movement and so on. And this process continues until the organizational state of the organism changes in such a way that the perception of the object doesn't cause the physical movement.

So the lion is intentionally moving its legs because the movements are a part of stalking the gazelle, which the lion is doing because it sees the gazelle and is hungry or playful and so on. We have now arrived at the second condition: a physical movement is intentional if it is part of an overall activity of the organism with respect to the perceived object.⁷

The third condition of the embodied theory is that the parts of the organism relevant to the extended causal relation between the organism and the perceived object must be healthy. I mean health here in the normal biological sense. A part of the body, such as the heart or a sense organ, is healthy if it is functioning properly in virtue of being in a stable relation with its surrounding micro-organisms.

Suppose the lion sees the gazelle, performs the physical movements of stalking it and then when it comes close enough to the gazelle to attack it, instead of attacking it or even toying with it, it tries to nuzzle with it and playfully bite it as if it were another lion. In this case the physical movements are a part of an extended activity of the organism with respect to the perceptual object. Nonetheless the physical movements aren't

⁷ This idea is defended in contemporary philosophy in two different contexts. Embodied cognition theorists argue that perception and action cannot be divorced from broader perception-action feedback loops. See Hurley 1998 and E. Thompson 2007. Neo-Aristotelians argue that actions are intentional because they are done in order to perform broader activities which constitute the life form of the agent; they thus defend a non-psychological, instrumental view of action. See Vogler 2002 and M. Thompson, unpublished.

intentional, for they aren't a part of the extended activity in the right way. They result from a malfunction within the lion and so they are an expression not of the organism but, as it were, of an alien force within it. The lion isn't stalking the gazelle because its relation to the gazelle has been transformed due to an abnormality within it. So in order for a physical movement to be intentional there must not be any such abnormality.

Two clarifications are necessary here. First, normal or healthy functioning of an organism is relative to a species. There are no a priori limits on what kinds of organisms there can be and therefore no limits on what it would be for them to be functioning well. For example, there could be an organism which sometimes stalks an animal and at other times treats the same animal as one of its own; and so its physical movements in both cases can be intentional. But from this fact it doesn't follow that any organism which engages in such behavior is acting intentionally, since this depends on what is healthy functioning for that organism.

Second, health is a matter of degree and so the intentionality of a physical movement is also a matter of degree. Suppose the lion stalks the gazelle and pounces on it, but because it has a weak heart it is not able to exert itself fully in order to take down its prey. The lion's pounce is still intentional, but it is less intentional and less full of life than the pounce of a vigorous lion at the height of its powers. This is clearer if we suppose the lion's health has deteriorated to such an extent that it is barely able to even stand up. It might still intentionally perform some movements which depend on bodily parts which are functioning relatively well; it might still move its head and slightly stand up. But given the deterioration of the body these movements are more the isolated activity of particular parts of the body and less the activity of the organism as a whole.

4.3. Objections

I will now consider two objections to the embodied theory of action. Here it is important to keep in mind that the embodied theory is a view of action and not of the ordinary concept of action. The former cannot be beholden to the latter: in light of theoretical reflection we may have to give up some ordinary beliefs about action.

The first objection is that the embodied theory doesn't pick out only agents. According to the theory, a physical movement is intentional if it is part of an overall

causal network with the environment in the right way. But then even bacteria would be agents and so would robots if they interacted with the environment in the right way. But certainly bacteria and robots aren't agents.

I agree bacteria aren't agents. Bacteria gain information and nutrients from their surroundings and are responsive to the environment. But they are not responsive to the environment in virtue of having sense organs which are causally connected to movements of the limbs and other parts of the body. In this sense they don't have bodies at all and that is why they are not agents. They have neither the internal organization nor the complex structured relation to the environment characteristic of agents. The difference between bacteria and agents is therefore one of complexity.

It is tempting to think it is so obvious that bacteria aren't agents that we don't need to look to biology to decide the issue. We should resist this temptation because being an agent is itself a biological issue: it is a matter of how an organism with a certain complex structure healthily interacts with the environment. If it is discovered that bacteria have primitive analogues to sense organs and limbs, I think it implies that bacteria are primitive agents and it would raise difficult questions such as whether an agent could be composed of less complex agents. We cannot a priori rule out or answer these kinds of questions.

Robots do have analogues of sense organs and limbs and so it is easier to think of them as agents. But if a robot just processes rules in performing even complicated tasks such as navigating through a building, it will only simulate being an agent.⁸ In order to really be an agent it is essential that the parts of the robot which enable the interactions with the environment be healthy. This doesn't mean simply that the parts must function well, as would be true even with the parts of a car. It means rather that the robot must be attuned to a vast array of life in its environment, and in this way have the defenses and the vulnerabilities of a healthy life form. Thus in order for a robot to be an agent, life itself has to be replicated in the robot so that it is a part of the overall structure of life in its environment. This involves a staggering amount of complexity, but it is in principle possible and so robots can be agents.

The second objection is as follows. The first-person perspective of the agent is the

⁸ See Dreyfus 1992.

perspective he has on the world in performing an action. This perspective is essential for actions, for it is in virtue of it that the agent is active. The embodied theory however identifies an action in terms of causal relations in the world and so it only involves the third-person perspective. The agent's perspective is left out.

I agree that the agent's perspective is essential for actions. But the embodied theory captures this perspective because it is implicit in an organism's having a body. The agent's perspective is not independent of the causal structure of the organism and its environment. Rather, it consists of the actions available to the organism in virtue of that structure. That an organism has a perspective at all means that it cannot interact with all features of the world at once. In this sense an omniscient and omnipotent being lacks a perspective altogether.

Suppose two lions are stalking a gazelle, and that one is to the left of the gazelle and the other is to its right. Then just in virtue of the difference in their positions with respect to the gazelle they have different perspectives on it: the lion on the left perceives certain aspects of the gazelle and the environment, and the lion on the right perceives different aspects. This means that the actions available to the two animals are different. The lion on the left might more easily attack the gazelle's legs or he can reach the gazelle quicker. And the lion on the right might be able to go right for the jugular or he can take a sharper angle to reach it. The position of each lion therefore affords it certain advantages and disadvantages: certain actions are possible and others aren't. In this way the lions' bodies orient them to the world in some ways at the cost of other ways, and this determines their perspective as agents.

The state of an organism's body also affects its perspective. Suppose the lion on the left hasn't eaten in three days; then in perceiving the gazelle it might immediately go into a stalking mode. Its extreme hunger might cause it to focus only on the gazelle and so its possible actions would be limited to those of stalking it. Suppose in contrast that the lion on the right is sick and doesn't have an appetite. The illness might restrict the lion's actions to such an extent that it might barely be able to react to the gazelle. The agent's perspective therefore isn't independent of basic physical phenomenon such as his spatial position or health. Rather, it is constituted by them.

The objector might now say that this isn't the sense of perspective he meant. The

perspective he meant is irreducibly first-personal. A purely causal description cannot suffice to capture the agent's perspective. For if it did, the agent's physical movement would be caused by the state of the world and not by the agent himself.

I think it is true that in an important sense an agent's perspective cannot be captured from the third-person perspective. But this is not because the agent's perspective is independent of his causal relations with the world. Rather it is because there is a difference between explaining the agent's perspective and occupying it. However extensively one describes the causal structure from without, one cannot thereby occupy the agent's perspective. For the only way to occupy it is to be the agent.

For example, even if one described completely the causal structure of the lion stalking the gazelle, one will not thereby have the perspective of the lion. The only way to do that is to be a part of the world in the way the lion is. One has to be in the position of the lion with the perceptual information it receives and with its bodily history; that is, one literally has to be in its place. The embodied theory therefore does not imply that the third-person perspective is the only true perspective on the world. When the causal relations necessary for being an agent obtain, the agent himself has a perspective on the world.

V. The Consciousness Claim

5.1. *The Argument*

The consciousness claim is that agents are phenomenally conscious. If an organism performs actions, it must be phenomenally conscious. There is a *prima facie* intuitive argument for this claim. Suppose a person picks up a piece of chocolate without being phenomenally conscious of any of its properties: he doesn't experience how it looks, feels and so on. Then to him it is as if there is no chocolate there at all and so he cannot intentionally pick it up. Without phenomenal consciousness, the agent cannot be active and is limited to how his body happens to move.

I think this argument is basically correct. But as stated it is not sufficient because it raises two questions. What exactly is the relation between consciousness and an agent's performing an intentional movement? And why does the consciousness have to be phenomenal and not just psychological? In order to answer these questions an elaboration

of the intuitive argument is needed.

Here is the elaboration of the argument:

1. *The affordance theory of action.* In performing an action the agent is conscious of the object towards which the action is directed as calling forth that action.
2. The affordance theory of action is possible only if the consciousness of the object is phenomenal.
3. So, agents are phenomenally conscious.

(1) states a necessary condition on how the agent must be conscious of an object in order to act with respect to it. (2) states that the condition can be met only if the consciousness is phenomenal. Together they imply the consciousness claim.

5.2. The Affordance Theory of Action

By consciousness I mean in this section a mental state which is intentionally directed to an object; the mental state is about that object. By an object calling forth an action I mean that it elicits that action. So an agent is conscious of an object as calling forth an action when he is intentionally directed to it by performing the action which it elicits. The affordance theory of action states that the agent is conscious in this way of the object towards which his action is directed.

Consider again the lion stalking the gazelle. Suppose that while performing the physical movements of stalking the lion is not conscious of the gazelle; it does not have a mental state intentionally directed towards the gazelle. Then the physical movements wouldn't be intentional because they wouldn't be guided by the lion's awareness of the gazelle. To the lion it would be as if there is no gazelle in its environment and so it would have no reason to move towards it. So in order to stalk the gazelle the lion has to be conscious of it.

A similar argument implies that the lion has to be also conscious of its movements and of a need it has which motivates the movements. Suppose the lion is conscious of the gazelle but is not conscious of its own movements towards it. Lets even suppose that the lion is conscious of getting closer to the gazelle, but is not conscious of its movements as what enable it to get closer. In this case the movements aren't intentional because the lion

does not consciously alter them in response to the gazelle. To the lion it will be as if it were following the gazelle independently of its own activity.

Suppose the lion hasn't eaten in many days and is hungry. Suppose further that the lion is conscious of the gazelle and of its leg movements but is not conscious of its hunger. In this case the lion's movements aren't intentional because consciousness of the gazelle isn't sufficient to motivate the lion to stalk it. If the lion isn't conscious of its hunger or of some other need which requires stalking the gazelle, being conscious of the gazelle would not motivate it to stalk it any more than being conscious of a rock or a tree would. Without the consciousness of its hunger, to the lion it will be as if it were moving towards the gazelle without a purpose.

Therefore in order to perform an action the agent has to be conscious of three things: the object to which the movement is directed, the physical movements and the need motivating the action. Being conscious of the three things however isn't sufficient to perform an action. One also has to be conscious of them in the right way.

Suppose the lion is conscious of its movements in the way it is conscious of objects in the world. It is conscious of them simply as happening and so as features of the world independent of it. In this case the lion is not intentionally performing the movements since it is only observing that its legs are moving and is not conscious of them as something it is doing. In order to be conscious of them as something it is doing it has to be conscious of them in relation to the broader activity it is performing, namely, that it is stalking the gazelle. Thus the lion must be conscious of its movements as directed towards the gazelle in order to attack it. In this way it is conscious of its movements in terms of the object towards which they are directed.

The way the lion is conscious of its movements implies that it is conscious of its hunger in terms of its orientation towards the environment. Suppose the lion were conscious of its hunger as a state within itself; just as it is conscious of things in the world through perception so too it is conscious of its hunger through a kind of inner perception. Then its movements would not be intentional because it would be observing the hunger causing the movements just as it would observe a causal interaction in the world. If the lion is conscious of the hunger in a detached way, it would also be conscious of its movements in a detached way.

Therefore in order to be conscious of the movements as a part of stalking, the lion must be conscious of its hunger as propelling it towards the object it is stalking. It must be conscious of the hunger not as a particular state within itself but as a general state of its body within the world. If there is no object to attack, the lion is conscious of its hunger in terms of being lethargic, of not being playful, of focusing on procuring food and so on. And if there is an object to attack, the lion is conscious of its hunger in terms of focusing all its activities on that object, and is conscious of its movements as a part of that focused activity. The lion is thus conscious of its hunger in terms of how it is orientated towards the world.

The way the lion is conscious of its hunger implies that it is not conscious of the gazelle as something independent of its needs. Suppose the lion is conscious of the gazelle only in terms of its properties such as its shape, size, position in the environment and so on. That is, suppose the lion is conscious of the gazelle in terms of objective properties it has independently of the lion; it is conscious of how the gazelle really is. To be conscious of the gazelle in this way is to aware of it in a disinterested way, as if the lion is standing back and just observing it. Yet to be conscious of its hunger is to be oriented towards the world in order to fulfill its needs and so to be aware of the gazelle in terms of its own interests. Stalking the gazelle and being conscious of it in terms of its objective properties are thus incompatible. The lion would then be moving its legs in order to satisfy its hunger even as it is conscious of the gazelle as if it were not hungry.

In stalking the gazelle the lion therefore must be conscious of it in terms of its own needs. In particular, it must be conscious of it as food, as something which will alleviate its hunger. Thus in stalking the gazelle the lion is conscious of it not as a gazelle but simply as something to eat. Its hunger propels its movements to an object which it is conscious of as eatable.

What is it for the lion to be conscious of the gazelle as something to eat? It means that the properties of the gazelle which are not pertinent to the lion's stalking it are irrelevant to the lion. Or, to put the same point differently, the lion is conscious of the properties of the gazelle only in terms of the actions it is able to perform in order to attack it. If the gazelle is large, the lion is conscious of it in terms of avoiding attacking it from the front. If the gazelle is far away, the lion is conscious of it in terms of the space it has

to traverse in order to pounce on it. If the gazelle has a broken leg, the lion is conscious of it as a weak spot to attack. For the lion in the midst of stalking the gazelle is, as it were, anonymous: it has no features beyond a thing to attack in a certain way.

A similar point is true of the lion's consciousness of the surrounding environment. Just as the lion is conscious of the gazelle in terms of actions necessary to attack it, so too it is conscious of the environment in terms of actions which aid or hinder its attacking the gazelle. If there are rocks between the lion and the gazelle, the lion is conscious of them as things to avoid in order to reach the gazelle or as things to hide behind in order to get closer to it. If there are other animals near the gazelle, the lion is conscious of them insofar as they might interfere with its stalking the gazelle. The lion therefore is conscious of the other entities in the environment only in relation to the thing to be eaten.

The lion when stalking is thus conscious of the gazelle not as one object among others in the environment. Rather it is conscious of it as the primary object of attraction in terms of which neighboring things are characterized and towards which it is drawn. The lion is therefore conscious of the gazelle as calling forth the stalking movements. It is conscious of the gazelle not as a recalcitrant object it has to act upon but as a magnetic space in the environment which elicits movements towards it.⁹

The lion's consciousness of the gazelle, its movements and its hunger are therefore united in the following way. It is conscious of the gazelle in terms of the movements it performs towards it. It is conscious of those movements as directed towards the gazelle in order to attack it. And it is conscious of its hunger in terms of being oriented to the gazelle as an object to attack. Thus when stalking the lion's consciousness of the gazelle and of itself are inseparably interwoven into a consciousness of acting in the world.

5.3. Action and Phenomenal Consciousness

In defending the affordance theory of action I was neutral about whether the consciousness in question is psychological or phenomenal. I will now argue that it has to be phenomenal, for only it captures the way the agent's consciousness of an object is related to his performing actions directed towards it.

⁹ For a classic statement of this view, see Gibson 1986.

Suppose the lion when stalking is psychologically conscious of the gazelle. This means that the lion has an internal mental state with the following properties. It is caused by the gazelle and thereby represents properties such as its size, position, color and so on. It is causally related to other mental states such as the desire to attack the gazelle, knowledge that attacking it requires certain movements and so on. And together with the other mental states it causes the stalking behavior. Thus for the lion to be psychologically conscious of the gazelle is for it to have an internal state with the appropriate causal properties.

Thus if the lion is psychologically conscious of the gazelle, it doesn't imply that it is actually performing the movements of stalking the gazelle. It only implies that the lion is in an internal state which is apt to cause stalking behavior if it has certain other mental states. Suppose there were two lions which gained the exact same information from the gazelle and so were in the same internal state with respect to it. Suppose further that one of the lions is hungry and the other is not, and so one lion stalks the gazelle and the other does not. In this case even though the actions of the two lions are different, their consciousness of the gazelle will be identical because they are in the same internal state with respect to it. The consciousness of the gazelle is thereby independent of whether the lion is active or passive towards it.

This however is not possible if the lion is conscious of the gazelle as calling forth the actions directed to it. For then the lion's consciousness of the gazelle is essentially related to its actually performing those actions; there cannot be former without the latter. The way the lion is conscious of the gazelle when it is stalking cannot be factored into a state which can exist whether or not the lion is stalking the gazelle. For being conscious of the gazelle as calling forth the actions captures the distinctly active way in which the lion is conscious of it.

One might object that this difference between consciousness as calling forth actions and psychological consciousness exists only if the latter is understood in too narrow of way. Instead of thinking of the lion's psychological consciousness of the gazelle as one internal state causally related to other internal states, we should think of it as the overall internal state of the lion. So if one lion is hungry and another lion is not, their psychological consciousness of the gazelle will not be identical since their overall

internal states will be different. This implies that the lion's psychological consciousness of the gazelle is not only apt to cause stalking but that it will in fact cause it.

This interpretation of psychological consciousness however still faces the same problem. For it is still possible that something could be in the overall internal state the lion is in without actually performing the actions of stalking. Suppose that there is a brain in a vat counterpart to the lion in the world such that the brain in the vat is physically identical to the brain of the lion in the world. This means that the brain in the vat lion and the actual lion will be in the same overall internal state and so be identical in terms of being psychologically conscious of the gazelle. But by hypothesis the brain in the vat lion cannot perform actions directed towards the gazelle since it is just a brain in a vat; it has no body to perform the actions with. So once again the lion's consciousness of the gazelle would be factored in a state which can exist whether or not the lion is actually stalking.

The lion's consciousness of the gazelle while stalking therefore cannot be psychological consciousness. Since the latter is defined as an internal state only causally related to actions, it is always in principle possible for the lion to be conscious in this sense without stalking. But it is impossible for the lion to be conscious of the gazelle as calling forth actions without stalking it.

Now suppose that the lion when stalking is phenomenally conscious of the gazelle. This means that the lion is conscious of the gazelle in terms of feeling a certain way; there is something it is like for the lion to be conscious of the gazelle. What is that feeling like?¹⁰

If we put ourselves in the place of the lion, we can imagine that it feels something like the following. Before it is conscious of the gazelle as something to stalk it might be conscious of it as just one more object in the environment. It phenomenally experiences the brownness of the gazelle and its white streaks against the background of, say, a lush green field or a watering hole with sparkling blue water. Even upon just seeing the gazelle the lion might be more alert and focused on it, and so might phenomenally experience the gazelle more vividly than it experiences the neighboring objects. But until

¹⁰ One might worry that perhaps we can't answer this question because we are not lions. But this is besides the point since the pertinent issue is the general point of what it is like for an agent to be conscious of an object towards which he is acting. If it helps, one can think instead of a human situation which is more familiar; say, what it is like for a person to be conscious of a pen he is reaching for in order to write with it.

it starts its stalking movements it experiences the gazelle as part of the open expanse of the environment in which any number of actions are possible.

However once the lion starts its stalking movements its experience of the environment and the gazelle changes. It now experiences the environment not as an open expanse but as structured around its activities aimed at the gazelle. What it is like to be conscious of the environment is that it is the space which separates it from the gazelle and so is the space which has to be navigated in a certain way in order to reach its prey at the other end. There is the gazelle at the one end and itself on the other end and it is conscious of all the other objects as a part of an indeterminate, open ended domain of everything else.

This doesn't mean that the lion doesn't experience differences between other objects in the environment; between say, the trees, stones, other gazelles and lions. Rather it means that it experiences those differences not as properties of clearly defined objects but as different properties of the background environment itself. It is conscious of the trees as brown, green areas to avoid and of the stones as hurdles in the environment to jump over. Similarly, it is conscious of the other gazelles, if there are any, in terms of having to be less focused on any one gazelle because there is an abundance of food in its surroundings. Thus it experiences the different objects in terms of the multifarious features of the background itself.

When the lion starts stalking its consciousness of the gazelle also changes. Whereas before beginning to stalk it might be phenomenally conscious of the gazelle as a determine brown entity, once it is in the midst of stalking it does not experience the gazelle as a determinate object at all. Instead it is conscious of it as a brown spatial region beyond its reach and towards which it is being pulled. Thus what it is like for the lion to be conscious of the gazelle is that there is something over there towards which a continuation of movements is necessary; it is what is at the end of this path. The lion therefore isn't phenomenally conscious of the gazelle in a way prior to its actions and which thereby causes them. Rather it is phenomenally conscious of the gazelle in feeling that its actions are directed towards it.

The lion's phenomenal consciousness of the gazelle it is stalking is therefore exactly what the affordance theory dictates. The lion is phenomenally conscious of the

gazelle as calling forth the actions it is performing towards it. It cannot be phenomenally conscious of the gazelle in this way if it isn't actually performing the actions of stalking, for what it is like for it to be conscious of the gazelle is that it feels its actions directed to it as to their conclusion.

One might now object that the problem I raised for psychological consciousness applies to phenomenal consciousness as well. For isn't it possible for a brain in a vat lion to be phenomenally identical to the actual lion? If so, that means phenomenal consciousness also isn't essentially related to the lion's actions.

I think it is not possible for a brain in a vat lion to be phenomenally identical to the actual lion. For the way the lion is phenomenally conscious when it is performing an action is that it feels active through the movements it is performing. The further it is able to go in the process of stalking the more active it feels, for it is the natural expression of its body's ability to be in synch with the environment. The movements therefore aren't incidental to its phenomenal consciousness, but its central feature. If the lion were simply sitting, it wouldn't be able to experience what it is like to be stalking anymore than it can actually perform the stalking movements while sitting. Thus even if the brain in the vat lion can be phenomenally conscious in some way, it cannot be conscious the way the actual lion is when it is stalking.

This point is obscured if phenomenal consciousness is identified just with what it is like to have an experience. Suppose the lion is sitting and gazing at a green pasture in front of it. Then in seeing the green grass there is something it is like for the lion; that experience has a phenomenal quality. But it is in principle possible for any experience to be illusory. Even though it phenomenally seems to the lion as if the grass is green, there might not be any grass there or it might not be green. In this sense what it is like to have an experience can seem independent of how the world is and even in fact of whether one has a body.

Phenomenal consciousness in the most basic sense, however, consists not of what it is like to have an experience but rather what it is like to be an organism. It involves what it is like to occupy a spatial region in a certain way, depending on whether the organism is upright, has a tail, lives in water and so on. And it also involves what it is like to move through space in ways characteristic of its body, such as whether it can fly, run,

swim and so on. Phenomenal consciousness in this sense therefore cannot be independent of the body and the environment, since it is the consciousness of being in the world a certain way. This form of phenomenal consciousness is so basic that in our normal thinking we do not identify it the way we identify the phenomenal properties of experiences which stand out. But it is essential for an organism to be an agent because it is the mode in which it feels itself acting upon and being acted upon by objects in the environment.

VI. Conclusion

In this paper I have argued that a zombie world is inconceivable. A physical duplicate of the actual world contains agents because an agent is a mobile, physical organism which is enmeshed with the environment in ways characteristic of its healthy functioning. And an agent must be phenomenally conscious because otherwise he cannot be actively directed towards objects he interacts with.

It is generally assumed that to deny the conceivability of a zombie world is to be a reductivist about consciousness and to deny that it is a mystery. My argument implies that this assumption is false. The zombie world is inconceivable not because we know *how* consciousness is a part of the physical world. We are now only at the very beginning stages of acquiring this knowledge and consciousness is certainly a mystery in that sense. Rather it is inconceivable because we know *that* it is a part of the physical world given its essential role in an organism's physical actions. The physical world, actions and consciousness are all essentially related to each other. Making progress on the mystery of consciousness requires better understanding these relations.

WORKS CITED

- Anscombe, G. E. M. 2000. *Intention*. Harvard University Press.
- Block, N. and Stalnaker, R. 1999. Conceptual Analysis and the Explanatory Gap. *The Philosophical Review* January, 1999
- Dennett, D. 1991. *Consciousness Explained*. Little-Brown.
- Dretske, F. 1995. *Naturalizing the Mind*. MIT Press.
- Dreyfus, H. 1992. *What Computers Still Can't Do*. MIT Press.
- Chalmers, D. 1996. *The Conscious Mind*. Oxford University Press.
- Chalmers, D. 2002. Consciousness and Its Place in Nature.
In Chalmers, D. ed. *Philosophy of Mind*, Oxford University Press.
- Chalmers, D. 2004. The Representational Character of Experience.
In Leiter, B. ed. *The Future of Philosophy*, Oxford University Press.
- Chalmers, D. 2005. Perception and the Fall from Eden.
In Gendler and Hawthorne eds. *Perceptual Experience*, Oxford University Press.
- Gibson, J. J. 1986. *The Ecological Approach to Visual Perception*. Lawrence Erlbaum.
- Harman, G. 1990. The Intrinsic Quality of Experience. *Philosophical Perspectives* 4.
- Hill, C. 1997. Imaginability, conceivability, possibility and the mind-body problem.
Philosophical Studies 87.
- Hurley, S. 1998. *Consciousness in Action*. Harvard University Press.
- Levine, J. 1983. Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly* 64.
- Loar, B. 1997. Phenomenal States. *Philosophical Perspectives* 4.
- Thompson, E. 2007. *Mind in Life*. Harvard University Press.
- Thompson, M. unpublished. Naïve Action Theory. <http://www.pitt.edu/~mthomps/>
- Vogler, C. 2002. *Reasonably Vicious*. Harvard University Press.