A Phenomenological Contribution to Analytic Philosophy of Mind

An Alternative Explanation of Consciousness

Signe Ulimoen Øverli

Master Thesis of Philosophy
Faculty of Humanities
Department of Philosophy, Classics, History of Art and Ideas

UNIVERSITY OF OSLO

12.05.2014

A Phenomenological Contribution to Analytic Philosophy of Mind

An Alternative Explanation of Consciousness

Signe Ulimoen Øverli

Master thesis of Philosophy

Faculty of Humanities
Department of Philosophy, Classics, History of Art and Ideas

UNIVERSITY OF OSLO

2014

A Phenomenological Contribution to Analytic Philosophy of Mind -

An Alternative Explanation of Consciousness

Signe Ulimoen Øverli

http://www.duo.uio.no/

Trykk: Reprosentralen, Universitetet i Oslo

Abstract

This thesis is an investigation of possible solutions to the Hard Problem in philosophy of mind. The problem describes the difficulty of explaining the abstract, experiential content of consciousness within the framework of empirical science. I here refer to the Hard Problem as it was put forth by David Chalmers (1995).

The main purpose in this thesis is to investigate how a phenomenological understanding of consciousness can be used to supplement the effort to find an explanation to the relationship brain and mind.

In order to find a plausible way in which this can be done, I discuss some different explanatory models for how consciousness can be understood and explained. Here I look at the difference between causal and constitutive types of explanations, and conclude that a plausible way to answer the Hard Problem is to find a constitutive explanation for how consciousness is realized (Craver 2007).

With this in mind some of the different efforts to explain consciousness are explored, with a special emphasis on reductive theories, such as the physiclaistic mind/brain identity theory and functionalism. As there are some evident problems with reductive physicalism as it is found in for example the mind/brain identity theory, I focus especially on functionalism and the way to explain consciousness proposed by this theory. In relation to this I also discuss the separation between role filler and role- functionalism (McLaughlin 2006).

Both these functionalistic theories get far in terms of explaining mental states, but there are still some difficulties connected to making sense of the experiential or qualitative aspects of consciousness. But by supplementing the theory with a phenomenological understanding of consciousness, as it is described in my reading of Alva Nöe's enacted approach, it is possible for functionalism to give a stronger account for the realization of consciousness.

Preface

This thesis is written as a contribution to the discussion around David Chalmers' Hard Problem in Philosophy of mind. But more importantly it is also an investigation of how the phenomenological understanding of consciousness can contribute as an explanatory factor in the effort to explain the evident gap between body and mind. I was motivated by the intuitive understanding of experiential consciousness found in the phenomenological writings by amongst others Edmund Husserl and Martin Heidegger, to find out how this way of describing consciousness could work within the framework of analytical philosophy. My thought was initially that analytic philosophy of mind could benefit from the profound understanding of experience and consciousness offered by phenomenology, and that this insight could be used as a starting point or a supplement in order to come up with a better explanation of the relationship between mental states and the physical world.

In order to find a plausible way for phenomenology to supplement analytical philosophy of mind I have also had to undertake a thorough investigation of this approach to consciousness. As this is an area in which I have had rather little experience up until now, writing this thesis has been an interesting and educational project.

In the end I hope to show that it is in some respects possible to get a better understanding of the relationship between experiential consciousness and the physical world.

Acknowledgements

I would like to thank my supervisor Ander Strand for all his valuable advice, and for helping me to pin point the direction of my thesis.
In addition I would like thank Arild Edsberg and Emma Skjølberg for sharing the experience and for seeing me through the whole project.

Table of Contents

1 Introduction

In this thesis I look at Chalmers' (1995) Hard Problem of philosophy, and discuss some possible solutions to it. In order to come up with an answer to the Hard Problem it is necessary to explain the relationship between conscious experience and the physical world/neural activity in the brain. The main topic is an investigation of how phenomenology, in the way it is described in Alva Nöe's enacted approach can be used by analytic philosophy of mind for the sake of creating a better, more satisfying answer to how we come to have the experienced and subjective qualities that are connected with consciousness. This is done through a discussion of what an adequate theory of consciousness should look like, where I look at both causal and constitutive models for explanation. I later move on to a discussion of the problems connected to some of the analytical explanations of consciousness, and propose some ways in which the phenomenological understanding of consciousness can be used to supplement these theories. In conclusion I propose that the phenomenological way to account for consciousness can be used as a supplement to functionalism in order for this theory to give an adequate explanation of the experiential consciousness.

1.1 Historical Background

In 1641 the French philosopher Rene Descartes published his "Meditations", describing how the world can be understood as split into two different substances.[1] One is the physical world, or what Descartes called res extensa, which is observable and open to empirical research through our senses. The other one, res cogitas is the immaterial or abstract substance of our consciousness; our experience of the world along with thoughts and feelings. These two different substances were supposed to be connected in the pineal gland, and this connection provided for the interaction between the physical body and the immaterial mind. Since the mind is perceived as something that is without any physical substance it is not available for standard physical research, or even submitted

[1] In addition to this Descartes also describes God as being a third substance (Descartes 2012)

to physical laws. In order to make inquiries, research and explain consciousness then, the only way of gaining knowledge is through introspection, as consciousness is out of reach for empirical methods of research.

This dualistic notion of two different kinds of substances, one for the physical and one for the mental, has left deep traces all the way up to modern philosophy in terms of how we understand the relation between body and mind. The theory, with its problematic dualistic aspect has also led to some severe problems for philosophy, within both ontological and epistemological areas, and the problem of dualism is still today a major obstacle for the philosophy of mind, as it leads to what we now know as the mind/body problem. The problem is that Descartes' solution to the mind body problem, with the pineal gland and some kind of consciousness spirits flowing through our veins in order to make the will of consciousness manifest in the physical body, is not on account with what we know to be true in the light of modern science. The relation that connects the mind and the body thus need to be explained differently, or else we will simply have to accept that our conscious minds have no connection to the physical world in which we live, and that our conscious selves must be understood as something that is closed within the body/brain, committed to only observing what is happening to our objective bodies like a person who is watching a film at a movie screen.

In order to solve the problems raised by this dualistic perspective of the relationship between body and mind, many possible solutions have been proposed. Explanations have focused on how it is possible for a non-physical entity like consciousness to somehow influence the physical body or brain. For those who do not think such an explanation will ever be possible, an alternative to the dualistic picture has been proposed, where the subjective consciousness are reduced to physical states. In this case consciousness can be explained by these physical states, rather than as something that differs from the physical world.

In modern philosophy of mind, the mind/body dualism of Descartes is to a large extent regarded as an insufficient explanation of consciousness. This means that most philosophers, with some exceptions, such as Zimmerman (2004) and Lowe (1992) wish to avoid this kind of fundamental distinction between body and mind. Nonetheless, for many theories in the field of philosophy of mind, dualism is still a problem it is hard to get away from, without having to deny the existence of the subjective and qualitative experiential mode of consciousness all together. None the less it is a strong tendency on the part of neuroscience, and even in philosophy of mind to figure out a way in which consciousness can be reduced to a physical instance, something that belongs to, and is the product of the physical world. The reason for this is that consciousness either have to be something that is possible to

explain within the physical framework modern science operates by today, or this framework will need to be altered extensively. This option does, as stated above, look like the best way to go about the problem of consciousness to most of the current scientists and philosophers in the field. As the other option, dualism, allows for there to exist something mystical and independent form normal physics, a notion that is rarely the interest, or position taken by scientifically minded people.

So the modern, scientific way of explaining consciousness is as something physical, most likely neurological action and signaling in the brain. Modern neuroscientists work strenuously to find neural entities or structures that correlate with qualitative experience in order to show how the brain is connected to, and responsible for consciousness. This view is supported in philosophy by reductive physicalism and other similar reductive theories, but as I will discuss in the course of this thesis, there are still some deep problems to the idea of reducing mental states down to physical states in the brain.

1.2 The Hard Problem

The problem that arise in relation to what I described above, have echoed through philosophy of mind, and the core of the problems has been pointed to by amongst others Thomas Nagel (1974). In the famous essay "What is it like to be a bat?" He points to the difficulty of explaining objectively, or even doing scientific/empirical research on the felt and experienced quality of consciousness. According to the standard understanding of consciousness found in analytical philosophy of mind, a defining aspect of conscious experience is that it is subjective, an inner state that is not susceptible to objective observation of any kind. Nagel's worry was that even though a scientist can find out all there is to know about the physical aspects of a bat, or any other living being for that matter, one does still not get any insight into the experience of being a bat (or that other person etc.). How the bat experiences itself and the world around it remains hidden, and no matter how hard we look at it, even if we open up its head and gaze directly at its brain, will get us any closer to how it feels to fly or having sonar vision. This problem lies at the heart of the physical reduction of mental states, and thus it is essential also for neuroscience when they try to understand

and explain the foundation of consciousness, and why we come to have this experience of our lives.

The problem, and the way it was pointed out by Nagel, was later adopted and developed by David Chalmers (1995). In his article "Facing up to the Problem of Consciousness" he formulated what we now know as the Hard Problem of philosophy in mind. He did this by contrasting Nagel's problem of finding out what it is like to be something/someone in terms of experience with other problems tended to by neuroscience, such as finding out and describing the difference between wakefulness and sleep, and being conscious versus unconscious and other similar problems connected to neural activity. The difference is that questions of the latter kind are available to scientific inquiry and research because they are determinable by observing the brain. The matter of solving these kinds of problems is one of having the right kind of instruments, and an understanding of how these things should be measured. If these things are in order, the question should be answerable within the existing framework of modern science. In other words the question is perceived as possible to explain, and the only things that are missing are the details of the supposed explanation.

But this is not the case with the Hard Problem, according to the reasons I mentioned above. The problem of how consciousness should be explained, and how it relates to the body when it is understood as irreducible to physical substance (the mind/body problem) is not a question with an available physicalistic or scientific answer, because it is not clear exactly how consciousness can be reduced to a physical, or at least scientifically viable instance. But how then can we hope to explain the fact that we are conscious? For intuitively we do know that we experience the world, as we are always somehow related to, or intended towards it.

1.3 Consciousness

The word "consciousness" has many different meanings, depending on the context in which it is used, and the person who uses it. To be absolutely clear about the way "consciousness" is used here, I will say a little bit about how I understand the term.

Consciousness, as the term is used here, refers to the subjective and experiential

qualities of mental states, and the first person perspective we have on the world. In other words I use consciousness in more or less the same way as it was done by Nagel (1974) and Chalmers (1995) in the formulation of the Hard Problem, or what can be referred to as qualia. This way to understand and talk about consciousness should be separated from, and be understood differently than, a reference to the physical state of being conscious, as opposed to being unconscious (in a coma or dead for example). It should also be separated from the background states of experiential consciousness, or the mode of consciousness we are in. This second division is more complicated to get a hold on, but according to Chalmers (2000) there are probably many more or less normal mental states that can be fitted into this category, such as for example being tired, inattentive, euphoric and so on. These are all contexts that influence the way we experience bot ourselves and the situated environment we are in. These two kinds of "consciousness" are both necessary and important factors for the qualitative aspect of consciousness that I refer to. To have such a subjective experiential consciousness it is crucial that we are conscious and not unconscious, and as we are conscious we are also necessarily in some kind of mood, or a given background states that dictates some of some of the way we experience the world. In order to determine and describe these underlying states of conscious experience it is often enough to look at the brain and it's neural activity. The state of being conscious has been found to correlate with a specific kind of neural activity, and if a doctor or scientist wants to determine whether a person is conscious or not it should be possible to determine this according to whether such activity is present or not. The same is true for other kinds of mental states or consciousness; their presence can often be determined by activity at a more or less special and restricted area of the brain. But as stressed by the Hard Problem, no such correlation is able to give us information about the subjective feeling associated with the neural activity.

The point is that consciousness is subjective rather than objective when we use the qualitative descriptions mentioned above. It is something we have access to through introspection, and personal reports about the way the world is perceived by the person in question. It is also expressed through behavior, as our acts tend to reflect cognitive contents. Another essential factor in relation to consciousness (as it is understood here) is that we, when we are conscious always are conscious of *something*. We are conscious of ourselves and of our surroundings, and are in addition able to perceive of other people/beings as conscious subjects as well. So it is safe to say that there does seem to be some kind of relation between the outside world and consciousness, even though it seems to be a difference between

consciousness and the kind of objective substance that can be measured and described in in a third person perspective. What separates the two is the way in which they are accessible for us in terms of research. When we want to explain a certain aspect of the world it is usually just a matter of empirical studies and experiments that can be performed within the framework of normal science and explained in accordance with physical laws. In the case of consciousness on the other hand, it is difficult to see how such objective studies can be done, because consciousness itself is understood as a subjective feeling.

The way the term consciousness is used in this paper is in other words as experience, the subjective or qualitative aspect of the mind, or what Chalmers calls qualia. When referring to consciousness in this way, it becomes clear that even though it is still possible to describe it as somehow connected to the brain, but it is difficult to see how we can say that consciousness is the same as a given brain state, or even more general neural activity. This is so, because detecting neural activity is not the same thing as detecting the felt experience a person has of its own self and surroundings

1.4 Why is Consciousness an Area of Interest for Philosophy?

Science and philosophy has long been connected, to the point where the two were seen as literally the same thing, all the way from ancient Greek philosophy in the antique and up to at least the time of Descartes. Science is still very much a part of philosophy today, but it seems like this relationship might not go the other way around. For instead of conducting philosophical reasoning and logical arguments, the method of modern natural science are prone to use empirical research and mathematical calculations to prove the truth of their hypothesis. The methods used by philosophy and science respectively are different, but the goal of the two are basically the same, to find out the truth about, and as such explain the world and the subjects living in it. But if the world can be made sense of by science and its physical laws, is it then any room left for philosophy. I think it is, as philosophy of science is still committed to clearing the grounds for natural science by questioning, discussing and optimizing the framework and methods used by natural science. But, even though philosophy does seem to be a useful tool for science, in terms of providing guidelines and modifications

to the way research is done and conclusions are made, in the case of cognitive science and philosophy of mind the relationship is even more on of mutual dependency.

The reason for this is that the concept of consciousness is difficult to grasp within the framework of normal natural science. The problem is that science usually is concerned with the kind of problems that are solvable through a normal use of scientific methods, but in the case of the rather abstract understanding of consciousness it seems like a different approach is needed in order to overcome the obstacles associated with the problem of the relation between brain and mind (i.e. the Hard Problem). What is needed is to determine a way in which consciousness can be understood as a physical entity, or at least a kind of entity that can be explained within a scientific framework (biology, psychology etc.). Here I think it is up to philosophy of mind to give a clear formulation of how consciousness can be understood in order for it to be put into an already existing scientific explanatory model, or alternatively provide a new model by which consciousness can be explained. The point is that, for now at least, neuroscience only has the appropriate tools to deal with research on the brain, and as such it is possible for scientists in the neuroscientific field to find out a lot about how the brain and the mind are correlated. What is still needed is a way for these correlations to be put into a system able to give an actual explanation of the correlation, but due to the intrinsic characterization of subjective and experiential consciousness it is not really possible to give an objective scientific description of this. The intrinsic characterization of mental stats does here refer to the notion that the experiential aspects of consciousness are understood to be hidden from public view. In other words these states are understood to be subjective, and as such inaccessible for scientifically guided empirical studies. So philosophy will have to do the job of giving scientific sense to the concept of consciousness. This is typically done by reducing consciousness to something that is within the reach of a scientific system, and these types of explanatory models will be the focal point of my thesis. But other suggestions have also been made, for example it is imaginable that the framework of physical laws can be altered or extended in order for consciousness to fit in (Chalmers
2006), or it is possible to hang on to the more mystical concept of dualism. One of the more extreme solutions to the problem is found in the skeptic views of Daniel Dennett (1991) and Paul Church land (1981), who denies the existence of consciousness all together due to its unscientific qualities.

1.5 Phenomenology

The necessary link between philosophy and science when it comes to providing an adequate explanation of consciousness, has been an inspiration for writing this thesis, and a lot of inspiration comes from phenomenology and the phenomenological understanding of consciousness. Phenomenology, in its classical form as it is found in the works of amongst others Edmund Husserl (2012) and Martin Heidegger (2000,
2007), is built on what might be called an unscientific ground, and is not constrained by scientific method. Rather this philosophical direction embraces introspection and first person studies as a way to gain crucial information about what it is like to be a subject living in an objective world. One of the goals of phenomenology is to make sense of both the world and consciousness through a skilled form of introspection, where the structures of how we access the world is discovered, and is as such a project that includes both ontological and epistemological investigations. The starting point for a phenomenological theory/description is the human being as a living subject in a given environment. The question of whether we are conscious or not does not even arise, and neither does the question of whether the world is real (as opposed to a mere illusion (Hume 2010, Leibniz 2001)). In phenomenology these things are taken for granted as they are both there, given by the way we exist. Instead of focusing on the world, and whether it is really there, phenomenology is occupied with the way it (the world) is manifested in our minds. The fact that we are conscious beings are crucial for this manifestation, had we not been conscious there would not be a world either, at least not for us. The notion that there is a world and a mind/subject that relates to it thereby stands undisputed. What is crucial for phenomenology is rather the way in which the world is given, or shows itself for us. As such, the theory is deeply intertwined with consciousness and psychology (of the introspective kind).

The aim of phenomenology is to make sense of, and explain what consciousness is, and what it must be like in order for the world to appear for us the way it does, and both Husserl and Heidegger (along with Sartre and Merlau-Ponty) offer extensive descriptions of the mind and how it relates to the world. In order to get an as clear as possible description of consciousness, the first step of phenomenology is to bracket the world, in a way that takes the

world in itself out of the picture[2]. In this way only the most important thing is left, consciousness, or the way the world is perceived by the observing subject. This "bracketing" is referred to as a phenomenological or eidetic reduction, and is meant to be a way for the phenomenological thinker to come closer to consciousness in itself instead of the world in itself. The meaning of the word eidetic is something similar to "see things clearly" or "seeing the clear picture", and this refers to the notion that only by taking the world in itself out of the picture do we get a clear insight into the ways of the mind. This is partly due to the phenomenal character of consciousness, or the fact that consciousness is always directed towards something in an intentional manner. In other words, consciousness has a phenomenal character (hence the name of the philosophical direction), and is always intended towards something, be that the world, a memory or another person. So the point of the reduction is to see the phenomena as they occur for the mind, to say it with Husserl (2012); consciousness is a transcendental instance, as it in a manner transcends the world.

The crucial point to take away from this is that only by viewing consciousness on its own, without concern for its physical relation to the world (the Hard Problem) can we hope to really understand it. The phenomenological description is, as stated earlier, concerned with the way the world is presented in our minds, and the phenomenology of especially Husserl and Heidegger describes many crucial factors of the way we relate to the world that I will not go further into here. The main conclusion is that both the world and the mind/consciousness seem to be given, or constituted by the way the two relate to each other. Had we not been conscious, no world would be constituted, or show up for us, and had there been no world (including our own and other beings physical bodies) for us to be conscious of there would be no conscious- ness. A phenomenological theory does in other words offer a thorough description of consciousness (in itself), and as such gives us a way in which to understand it. And clearly a good explanation of consciousness should start with a careful understanding of the phenomena to be explained.

[2] Epoche (Husserl 2012)

1.6 Phenomenology as a Supplement to Functionalism

From the discussion of phenomenology above, we see that the "mystical" notion of consciousness is treated somewhat differently by phenomenology than it is by the more scientifically oriented tradition of analytic philosophy. My main intention for writing this thesis has been to investigate how the Hard Problem (Chalmers 1995) appears in the light of the Phenomenological understanding of consciousness. My thought has been that the understanding of phenomenological consciousness can be helpful for analytic philosophy and neuroscience when it comes to bridging the "explanatory gap" between the physical body and the abstract concept of mind. The main topic of this thesis is thereby an investigation of whether a phenomenological insight into consciousness can be used as a supplement to analytic philosophy in order to better the understanding of how consciousness relates to the physical world, and how this can be done.

In what way can phenomenology be used as a supplement? And which analytical theories of consciousness are prone for this kind of supplement? Are some theories more closely linked with phenomenology, and as such a better theory to supplement with the phenomenological insights? These are the kinds of questions I deal with in this thesis, and in order to arrive at conclusions in relation to these, I will also investigate some of the foundations of classical analytic theories of consciousness.

I start out by discussing two kinds of explanatory structures that can be used on consciousness, in order to explain how it is connected with the physical world. An explanation of this relationship can be understood either as causal, or as constitutive. In the former kind of explanation the answer to the Hard Problem will be characterized as describing a link of cause and effect between physical states in the brain and the experiential content of the mind. By investigating, and arguing for the use of causal explanations in modern science, I shed some light on the way this kind of explanatory model can be seen as relevant for the understanding of various phenomena. The notion of a causal explanation also lays the grounds for how we can understand and use a constitutive explanatory model, when we want to explain how consciousness is realized by an empirically available entity or structure.

This leads to a discussion about constitutive explanations, what they are, and how they are put to use. Here I will use work done by Carl Craver (2007) in order to illustrate both the method and purpose of this constitutive model for explanation. The point here is to get an explanation that does not appeal to a cause and effect relationship, but instead has as its goal to explain a phenomena by appealing to the underlying structure of the way it is given (realized). I will argue that such an explanation fits better with what is looked for in the Hard Problem, as it does not demand the same kind physical and temporal structure as a causal explanation.

After discussing the constitutive explanatory model, I will follow this up by looking at some different efforts to explain consciousness and the mind/body problem in such a constitutive manner. The first theory I focus on is the mind/body identity theory, which, strictly in line with physical reductionism, holds that a mental state can be identified with a given state in the brain. This kind of theory is closely related to neuroscience and cognitive research, as it assumes that it will be enough to identify the brain states that correlate with the mental states in order to explain it (as these two are essentially the same). My argument here will be that the mind/body identity theory fails to explain consciousness for very much the same reason as reductive physicalism in general, and, instead I will investigate whether functionalism is a better option in terms of finding an answer to the Hard Problem.

In my discussion of functionalism we meet two different interpretations of the theory, role filler-functionalism and role-functionalism. Both theories hold that mental states can be described as functional states, the relationship between sensory input, behavioral output and other mental states. In role-functionalism this relation itself (the functional state) is identified with consciousness, while in the case of role- filler functionalism, the functional states are used as designators for the underlying physical states that are taken to be the realizing factor for consciousness. But a problem for both is still that even the functional descriptions does not open up for an understanding of the intrinsic qualities of experiential consciousness as long as consciousness is understood in the classical way it is described by Chalmers (1995) in relation to the Hard Problem. These intrinsic qualities of consciousness can be seen as the subjective or private feeling of experience, and if this aspect cannot be accounted for by functionalism, it is unsuccessful in terms of explaining consciousness.

Through a thorough review of the phenomenology as it is interpreted by Alva Nöe in his enacted approach to consciousness, I will argue that both kinds of functionalism can be

possible to supplemented with the view of consciousness we meet here. The enacted theory is, much like phenomenology in general, concerned with the way in which the world shows up for us in our consciousness, and what the mind must be like in order for this to happen. In Nöe's modernized phenomenology, it is a point to give scientific, empirical proof for why we should perceive of the conscious mind in the way it is stipulated by phenomenology, and as such there is hope that the theory can be taken seriously also by scientifically minded people. The main point of the enacted theory is to show that even though consciousness might be reduced to things that can be explained more or less scientifically, it takes more than a brain to constitute consciousness. Following this, it is possible to understand the enacted theory as offering a way in which functional descriptions also can be used as denote the intrinsic, subjective experience that is consciousness.

With the phenomenological and enacted comprehension of consciousness, I hope to be able to show that with a wider perspective on the constitution of consciousness it is possible to supplement the analytical theory of functionalism, in order to get a better understanding of the relationship between body and mind. In my opinion the phenomenological understanding of consciousness can prove to be an important factor when it comes to bridging the explanatory gap between body and mind as it is formulated in the Hard Problem because it offers an deeper understanding of consciousness, that is closer to the way we intuitively relate to both it and the world. I will also propose an alternative reading of Nöe's enacted approach that can be seen as handling the Hard Problem by explaining it away. This reading presuppose a view of consciousness as something that is already both present and presented to us because we are able to perceive the world and act in it. As such, it is possible to say that a Hard Problem never arise, as consciousness is already explained and understood by virtue of this interactive relationship.

2 Causal and Constitutive Explanations

The notion that consciousness has a qualitative, subjective and inert aspects reflects how we experience ourselves and the world around us (qualia) presents some serious problems for modern neuroscience and philosophy of mind (Chalmers 1996, Nagel 1974). This has rendered debate amongst philosophers about how this aspect of conscious life should best be understood and explained. A starting point for most philosophers is the scientific understanding of the brain, and how the structure of this physical entity gives rise to conscious experience. Most of the cells that process the information made available by the senses are found to be in the brain, and most modern scientists and philosophers accepts that the brain plays a more or less (Nöe, phenomenological thinkers) crucial role for the formation of conscious content. This thought, that the brain somehow gives rise to consciousness is also the normally accepted materialist view of consciousness, a branch of philosophy holding that the ultimate nature of the mind is physical, and that there is no sharp contrast between physical matter and consciousness (Dennett 1991). The materialist view can also be seen as a contrasting answer to the dualistic thought of body and mind stemming all the way from Descartes, who made a strong distinction between the physical and extended body and the spiritual mind. This dualistic way of thinking survived for a long time in the realm of psychology and philosophy of mind, but has in modern times, to a large extent been abandoned in favor of the materialistic view or other suitable theories in order to transcend the ontological gap that materializes between the physical and the mental in dualistic thinking. Descartes dualism (2012) led to numerous problems in the way we think of and understand consciousness. His view of the strong division between the physical and the mental has to a large extent been refuted by later philosophers and scientist (Chalmers 1996) in such a way that we no longer have to imagine the world to be made up of two completely different and incommensurable kinds of "materials". How this is done is not a pressing matter here, and I will not discuss it further, but it is none the less important to keep in mind that there still are some traces of the Cartesian view to be found in the way we talk and think about consciousness.

These traces can be found in the way we understand and think about conscious- ness and its physical "counterpart", the brain, as two different types of phenomena (Yablo 1992). According to the materialist view in philosophy of mind there are no significant differences between physicality and conscious content, and one should not yield to talking about qualia as anything over and above the neurological signaling in the brain (Smart 1959). In spite of this there is, as stated above, still some dualistic traces to be found in the way the relationship between brain and consciousness is apprehended. The dualism is apparent in how mental phenomena on a regular basis are treated as something different than physical phenomena (Descartes 2012). The dualistic notion is apparent also in the materialistic view where conscious content is often treated as supervening on physical action in the brain, making it seem like what is supervening is not itself a physical substance (Smart 1959). This illustrates what lies at the heart of the Hard Problem (Chalmers 1995). The still apparent di- vide between the mental and the physical plays up the long standing problem of how something physical can give rise to something mental, and if and how this mental content in some way differs from the physical world. These are important questions, and strong arguments are needed to support any of the stands, but this will not be the problem developed here. Whether conscious content and experience can be proven to be something physical or something "over and above" the physical, the Hard Problem still illustrates the question about the relationship between the physical and the mental still needs further discussion in order to answer both if and how the physical structure of the brain can give rise to consciousness.

So the Hard Problem demonstrates the need for an answer to how conscious content can possibly stem from or be realized by the chemical processes and electric signals in the brain. There is after all a big difference between what can be objectively observed and tested for in terms of third person observation, and the way what is thus measured is experienced by the person to whom the brain belongs. There is in other words some kind of division between the physics of the brain and the first person experience this physics gives rise to, the qualitative consciousness. Thus the Hard Problem is a hard one, and an explanation for the sake of the understanding of the relationship between the two is needed.

The hard problem does in other words need an explanation of what the relationship is between the physical structure of the brain and the conscious experience it gives rise to. But what should an explanation addressing this question look like? And what enables an explanation to provide an acceptable answer? In science and especially in

philosophy of science there are many theories and hypotheses about what a good explanation should look like. In this chapter I will go through a few of them in order to clarify what an explanation of consciousness should contain, and what kind of explanatory fallacies should be avoided in the search for an understanding of how mental phenomena and conscious experience springs out of the brain.

2.1 Causal Explanations

Explanatory models are often based on causal and constitutive forms of explanations, one of them, or both according to the kind of explanation the question demands. So the first step in the process of finding a solution to the Hard Problem should be to determine whether the relationship between brain and mind is causal or constitutive, and the widely accepted theory today is that this relationship should be seen as constitutive[3]. There are of course many reasons for this, and on a whole it seems like the constitutive explanation fits better with how consciousness relates to the physical structures and signaling in the brain (Craver 2007). At the same time the difference and relationship between constitutive and causal models of explanations is not always fully understood and expressed. To get a clearer understanding of this I will take a look at what is meant by both causal and constitutive explanation, and take a look at how they differ from each other and how they can be used in separate ways to explain given phenomena.

Many explanations in science are answers to questions about the relationship between cause and effect, and these questions seek causal explanations. In standard modern science a scientist will probably have to answer questions of both a causal and a constitutive character, even though it is often (Craver 2007) thought that all scientific answers should be based on causal explanation (but, as we'll see later this might depend on the question in need of explanation). A causal explanation is the default answer given to a so called "answer seeking why-question" (Ylikoski 2013, van Fraassen 1980), questions about why things are the way they are, and the normal way to respond to this is by appealing to some kind of causal explanation, making clear that the reason why a thing is this way rather than

[3] 3 Apparent in the way philosophers of mind build their explanations, all the way from analytic and scientifically inspired thinkers such as Crick and Koch (2003) to the more phenomenologically inspired theories of Alva Nöe (2009) and Dan Zahavi (2008).

another is because something else caused it to be like that. When asked about what might be the cause of a given event such as (to use a popular example first provided by David Hume (2010)) the movements of billiard balls, it seems reasonable to reply that the momentum of the cue or optionally the other billiard ball caused the ball to move.

This type of explanation seems adequate enough as long as the cause (one billiard ball hitting the other) is grounded in provable scientific facts (the momentum/energy is transferred from one billiard ball to the other causing the second ball to move). However there are still some deeply rooted metaphysical problems connected to the causality and its adequacy as a legitimate ground for explanation. Some of the issues here stem as far back as David Hume's "An Enquiry concerning Human Understanding" from 1748 (2010), where he denies the possibility of proving the existence of causal relations, thus making causal explanations invalid from a metaphysical/analytic point of view. An answer supported by a causal form of explanation will according to Hume appeal to the inductive fallacy, where the only thing really lending support to the explanation is that the relation has been observed before. Turning back to the question about how to explain the movement of a billiard ball across the billiard table, there is an observed correlation between one ball hitting and other, seemingly causing it to move. According to Hume the only thing that is really a fact here is that every time a ball hits the other this one will also start to move, but to say that one is actually observing the one causing the other can never be really true according to Hume. To say that the movement of the ball that hits the other also causes it to move is to confuse correlation and causation, and to confuse the two would be to step into an inductive fallacy. This way there is nothing implicit about correlation, the fact that a given thing have a habit of happening in a specific order that makes it true that they stand in a causal relationship to each other, so analytic predictions cannot be made from this base alone.

Many attempts have been made to avoid this fallacy of course, but from a metaphysical point of view it seems like there might be no final solution to Hume's problem. The way science is done today none the less seem compatible with the use of the causal method when a (scientific) phenomenon is explained and causal relationships can be tested in different kinds of experiments. If an "effect" never occurs without the presence of a given cause (or causes), this would imply that the explained phenomenon or relationship is one of cause and effect. From a scientific point of view causal relationships can be revealed through experimenting, by checking which things are able to alter or cause other things/events

to happen. The same is true for many of the other "causal problems" of the inductive fallacy. The important thing is that for everything that happens, for all phenomena that exist there is some sort of explanation or reason (scientifically speaking) for its occurrence (Craver 2007). The notion of such a principle originates amongst others from the German mathematician and philosopher G.W Leibniz. His principle of sufficient reason states that nothing can happen without a reason, and that for every event there must exist a reason sufficient for that event to happen (Leibniz 1989). But this explanation or reason is not always readily available, and the job of the scientist is not only to understand these reasons, but also to make explanatory representations of the phenomena. A representative explanation represents the phenomenon by explaining why it is the way it is, and it is here we meet the problem of how an adequate explanation should look like. Much effort has been laid down, partly by philosophers of science to make explicit exactly what should be expected of an explanation to make it a good one, and how such criteria need to be fulfilled in order to make up a good, coherent explanation.

One such strategy for making out what scientific explanations should look like is found in Carl Hempel's (1965) Covering Law model (Craver 2007). This model makes scientific explanations that appeal to causality plausible by proving how an outcome can be expected (predicted) on the basis of physical laws. The explanation is set up as an argument with an if-then structure where basic physical laws serve as premises. The argument then shows how in certain conditions the arguments conclusion (the phenomenon to be explained) is to be expected on the basis of these physical laws. Thus Hempel (1965) makes explanation a matter of using deduction, and thus avoid the inductive fallacy. On the other hand there are, according to Craver (2007) and of course many others, severe problems about causality left unanswered by the covering law model. First it is no restrictions embedded in the theory that makes it possible to separate between real physical laws and mere accidents. As both can be empirically accounted for, it is crucial for the model of covering law that the explanatory relevant physical laws can be separated from mere correlations and generalizations. Another problem for covering law is that it does not provide a sufficient mechanism to deal with explanatory relevance. A successful model of explanation needs to be able to account for the factors that are actually relevant for the explanation of the phenomenon under discussion and which ones are not. This moment of critique has also been stated earlier by amongst others Wesley Salmon, who pointed the problem of relevance by showing the way Hempel's method could yield true explanations for false conclusions

(Salmon1984). To find the relevant factors that can genuinely be posed as causes, scientists can of course do some surveys or experiments to see what the actual explanation of a phenomenon is (van Fraassen 1980), but in a model that treats explanations the same as arguments irrelevant factors can possibly sneak in without any danger of altering the conclusion, and as such, Hempel is again prone to fall into the inductive fallacy. This is because the true premisses of the argument leads to consistent conclusions even with the addition of irrelevant extra arguments. In addition, the model of covering law, with its focus on how the phenomenon is to be expected (by deduction) on the background of the physical law and relevant circumstances is not able to account for the many circumstances in which all the relevant premises are present without yielding the predicted outcome.

Hempel's mission is to make scientific explanations indisputable by setting them up as logically consistent arguments, where the conclusion of the argument or the occurrence of the phenomenon is seen as a necessary consequence of the premisses. But, as shown by Craver (2007), the theory is not altogether coherent with the reality of science and how scientific explanations are actually made, an important factor being that scientific explanations seldom or never are understood as arguments. Bas van Fraassen (1980), for example, was aware of this and tried to solve some of the problems in Hempel's method by adopting a version of an "anti-realist" view of Hempel's explanatory model. He follows Hempel's way of setting explanations up as logical constructions, but he switches out Hempel's logical relation of a cause and effect with the relation between question and answer. The reason for this is that van Fraassen believes that Hempel's understanding of scientific explanation is a misunderstanding of how scientific explanations are made. To find scientific answers and make explanations based on empirical research is not a matter of describing a relationship between cause and effect; rather it should be understood as en effort in terms of finding an answer to the kind of answer seeking-why questions I mentioned earlier in reference to Ylikoski. The point is that scientific explanation is a process of creative construction of explanatory models, where the aim is to find answers to questions, rather than an effort to reduce the phenomenon to the premises on which it is based. So instead of answering if a given phenomenon is equal to some deductive argument van Fraassen focuses on the why form of the question, and he answers by appealing to the best avail- able reason/cause. By doing this, van Fraassen avoids some of the inductive fallacies that Hempel falls into. For example, through van Fraassen's approach it is possible to avoid some of the problems that

are connected with explanatory asymmetry and relevance, by doing an evaluation of the kind of context the questions are asked in relation to. This means that the explanatory relevant causes can be determined by looking at what the relevant cause is in the given context of the specific incident that is under research. This being a clearly pragmatic strategy, is unquestionably more on terms with how research and explanations are actually performed by scientists. A point that is crucial for Craver (2007) in his view of causal explanations.

So, according to both Ylikoski (2013) and van Fraassen (1980) there is a clear misunderstanding in Hempel's understanding of how the link between cause and effect should be understood. By appealing to a subjective and pragmatic understanding of this relation, with support from how scientific explanations actually are made (based on answering, not discovering), many of the inductive fallacies connected with causality can be avoided. In result, this view serves as a solution not only to some of the problems encountered by Hempel, but also to those pointed to by Hume. This more proper understanding of how science explains the world is also crucial for the explanatory theory developed by Craver (2007).

Further criteria for what makes up a good (causal) explanation other than Hempel's deductive nomological theory of explanation is in other words needed in order to give an adequate model or theory for what causal, scientific explanations should look like. The shortcomings in the covering law theory pointed out by Craver serves well to illustrate the kind of criteria a well formulated model of explanation should meet. A sufficient form of explanation is one that accommodates the lacks described above, and, according to Craver (2007) meets certain criteria in order to count as a good explanation of the given phenomenon. These constraints, or the necessary content of a causal explanation, can be summarized as (1) mere temporal sequences are not explanatory, i.e. the explanation needs to describe an internal relationship between cause and effect over and above the fact that one comes after the other. (2) Causes explain effects and not vice versa, the relationship between cause and effect is asymmetric. (3) Causally independent effects of a common cause do not explain one another. (4) Causally irrelevant phenomena are not explanatory, all the factors or phenomena in the explanation need to be relevant for that explanation or else they are not explanatory. (5) Causes need not make effects probable to explain them, improbable effects are still effects, and the explanation must allow them to take place (Craver 2007 p. 26).

A good causal explanation thus needs to be careful in terms of fulfilling these criteria, in order for the causal explanation to hold in a real explanatory way. In many

theories of causal explanation the notion of causality entails some kind of over- lap between the cause and effect, as in the examples given by Craver in "Explaining the brain" (2007). One of these are the theory of mark transmission (MT) proposed by amongst others Wesley Salmon, which incorporates a view of causality as a form of touching between the cause and effect in which the cause transmits some kind of mark on the effect. The point here is to look at the relevant particles or objects in the causal explanation as individual "world-lines" in a space-time diagram (a Minkowskian space-time diagram (Geroch 1978)). This goes to show how some processes have actual causal relevance, while others have not, by looking at how an impact can be made by one object, so that it changes the world-line of another. The difference is made between a process that successfully alters the world line of another beyond the point of interaction and those who merely show an impact on the place where the lines meet. This is surely a nice way to discriminate between a causally relevant process and a process that only seemingly makes any kind of causal difference. The transmission theory is not on the other hand able to accommodate and explain processes where there is no actual connection or "touching" between to world-lines. But causality in neuroscience (as described by Craver) does not entail this kind of touching as a necessary factor of causality. Many neurobiological cases do not demand any direct connection between cause and effect, such as inhibitory mechanisms that prevents or stops the occurrence of a phenomenon by the creation of a gap or disconnection in the neural structure of the brain (negative causal relation- ships). In other words, this theory also shows weakness in terms of giving accurate explanations, at least in the field of neurobiology. According to Craver (2007) the MT theory's explanatory relevance lies in placing the world within a causal frame- work, but even though the model might be successful at this it does not accurately discriminate between causal connections relevant for a given phenomenon and which factors are not, and as such the MT theory is not an adequate theory of explanation because it does not meet all the standards for accurate explanations as they are set up by Craver.

Another important weakness in the kind of theory we meet here is that it keeps making an appeal to the causal "link" between cause and effect. The theories support themselves by appealing to connections on the fundamental level, and as such end up with letting the same old "causal worry" present itself here. The "remedy" for the problem of causal relevance as it is proposed by Craver (2007) is to look at science from a somewhat different angle, and appeal to how modern science is actually done. The strength of the appeal to and

need for strong causal connections on the fundamental level depends on whether we treat science as something that should be reduced to the most fundamental level of physics or can science be multileveled and split up in a hierarchical manner to give each level a fundamentally constitutive function. In Carl Craver's "explaining the brain" (2007) we meet a fundamentally anti-fundamentalist view, where there is no need to ground the "higher level" sciences in more fundamental structures if the purpose is to explain a phenomenon or mechanism on the stated level. This makes it possible to treat every level as self-sufficient with its own explanatory relevance, avoiding reductionist problems such as uncertainties about causal relationships on lower levels.

The notion of multiple levels in science is founded on the notion of mechanisms working on different levels of interest for scientific research. Although all mechanisms are built up by institutions on a lower level, the working of the specific mechanism is seen as working on a distinct level. One can in other words make a demarcation be- tween levels, by letting the working mechanism stand as an indicator of the particular level. In this way explaining how the mechanism work on a given level to constitute a phenomenon, does not demand any further confirmation or explanation on a lower level, as long as it is sufficient to explain the phenomenon of interest.

The mechanism itself can be described as an activity that to some extent can stand as a causal explanation for the occurrence of a given phenomenon, by being sufficient and necessary for that phenomenon (Craver 2007). This implies that without the underlying mechanism the phenomenon would not occur.

According to Craver, it is thus necessary to look at the phenomena to be explained in order to decide what a good explanation for it should look like. In many higher-level sciences such as neurobiology the phenomena to be explained, the explanandum, are mechanisms that function as an underlying reason for the occurrence of the phenomenon. By looking at the explananda as a mechanism, and by adopting certain normative rules as the ones that are pointed out above (1-5) it is possible to correctly identify causally relevant factors within a mechanism. With this under- standing of the causal relevance in mechanistic explanations Craver lays the ground for a different kind of explanation, making out how explanations also can reveal constitutive relevance. And even in the case of Leibniz and his principle of sufficient reason that we met above, it is not clear that the reason for the occurrence of an event have to be explained in a causal manner, and so there is also room for constitutive types

of explanations, as long as it can give sufficient reason for the occurrence of a given phenomenon or event.

2.2 Constitutive Explanations

So, from what is stated above there are two kinds of scientific explanations, and they might be of a causal or a constitutive character, depending of what kind of, and in what way a phenomenon demands an explanation. A constitutive explanation, as opposed to a causal explanation is explanatory relevant in terms of explaining the underlying mechanism of the phenomenon. The idea of the constitutive explanation is that phenomena in many cases can be explained without making an appeal to the kind of temporal structure with cause following effect through time (appearing after one another in a space-time diagram), but the difference between the two types of explanations is unclear and easily confused as they are both trying to explain some- what similar scientific problems. The difference between the two ways of explaining scientific phenomena might be better understood by illustrating what the two have in common and where they separate in terms of giving different kind of explanations to different phenomena. Firstly both of these models are aimed at making sense of, and explain the world and everything in it in a scientifically consistent way, they differ significantly in the way they address problems, and in what kind of things they explain.

A causal explanation describes a temporal sequence with causes and effects following each other in a fixed order, from cause to effect in an asymmetric relationship, meaning that the effect can never explain the cause. We can very often see this as an historical account, where cause and effect follow each other through time to get us to the event we now observe. The cause need also be something that is necessary for the effect to occur. A causal explanation as such will be an explanation of how one phenomenon works to change or create an effect, i.e. How these two, the cause and the effect relate to one another. A causal explanation will in other words describe the behavior of a mechanism, stating the kinds of conditions (the cause) that are necessary for the effect. For short one can say that the explanandum in a causal explanation, the object the causal explanation is to make sense of, is the relationship between cause and effect. Causal explanations (if they are successful) do seem like a good way to go about explaining the world, but often scientific explanations are

not only causal. Many of the explanations we find in the natural sciences are based on the constitutive model of explanation, and such explanations are intrinsically different than the ones supported by causality. Still some of these differences are unclear, and even amongst philosophers who are devoted to theories of explanation sometimes have difficulties with discriminating between the two (Ylikoski 2013). It is therefore crucial to take a closer look the individual and shared characteristics to understand how the two are related and separated.

The constitutive explanations, unlike causal ones are not intended upon explaining the behavior of a mechanism; rather its explanandum can be understood as causal capacities. The causal capacities are different dispositions, tendencies etc. Of an entity, that describes what would happen to the entity in any given causal setting. A causal explanation would then be one that explains why the disposition are realized, while the constitutive explanation is describing how these dispositions or capacities work, and what they are like. Unlike the causal explanation, the constitutive model does not explain temporal sequences or how cause and effect are related to each other through the span of time (as mentioned above the cause and effect does not occur on different points in the space-time diagram), rather the phenomenon that is sup- ported by the constitutive explanation is simultaneous to the constitutive mechanism. This simultaneity is apparent, or "caused" due to the explananda of constitutive explanations, which is the structure of the causal parts of a given mechanism. The constitutive explanation is thus, as is the case with causality, an asymmetric relation, but here the asymmetry points to how the constitutional system explain the causal capacities of this system, and not the other way around (the causal capacities do not constitute the causal parts and their organization) (Ylikoski 2013).

The relation between the causal parts of a mechanism and the way these are constituted it is aimed at making sense of seems to give and impression of constitution as somehow building on causal relationships within the mechanism, but the relevance of the explanations are still different as they essentially aim at explaining different things. Here are some tricks to understand how they differ, and the different kinds of questions they explain.

The normal way to separate between causal and constitutive explanations is to distinguish between etiological and constitutive explanations. Where the causal explanation has the relationship between cause and effect as its explanandum, the constitutive explanations aim at explaining the causal capacity of a system (Ylikoski

2013). A given system (for example a neuronal structure in the brain) can have many and varying causal dispositions or possibilities that can be triggered in the right sort of circumstances, and the goal of the constitutive explanation is to make sense of the causal capacities (the dispositions or properties mentioned above) and how they work as they do to make out the overall structure of the whole. So even though there are many similarities between causal and constitutive explanations they differ clearly from each other in terms of the explananda, or what they are actually trying to explain, and the explanatory relevance of the two models have different dependencies in terms of what make them into adequate explanations. To get a better grip of this it is necessary to understand what is meant by a causal capacity. The notion of a causal capacity illustrates what could be exacted to happen to an entity if it were put in a given causal setting, or in other words what would happen to an entity in a certain triggering or enabling condition. Background conditions such as temperature etc. Also play a role here, but this is a question of pragmatic relevance and can be determined by empiric inquiry. So the causal capacities of a system can behave in different ways according to the causal setting, and the job of the constitutive explanation is to explain why, and in what cases a causal structure gives (rise to) certain kinds of properties. One of the explanatory models, the causal one has as its object to explain events, the behavior of entities to give a reason for how something can be the cause of the other, while the other, the constitutive explanation wants to make out and explain what the necessary properties of a system (or even an individual causal capacity) should be in order to produce the causal event or process. To illustrate this difference and to shed some light over the way different kinds of explananda will make claims to different explanatory models, Ylikoski (2013) exemplifies three different ways to inquire about the fragility of a glass. First we can ask how a glass has become fragile (enough to break), we can ask what makes the glass fragile, and we can ask about why the glass broke. Of these three questions the first and the third demands causal explanations. They are inquiries about an historical process leading eventually to the fragility and breaking of a glass, and as such they fit in with the view that causality deals with processes and events. The answers to questions should in other words describe the process that has led the glass to become fragile and break, with the latter question (3) describing the triggering event that actually caused the glass to break. The second question on the other hand asks for a different kind of explanation. When asking about what makes the glass fragile, the point is not no gain an understanding of the historical process leading the glass to break, but rather what intrinsic qualities makes the

glass prone to break in the first place. The thing that needs answering in this question is what gives the glass the causal capacity to break, and such an explanation will be a constitutive one. This example, given by Ylikoski (2013) demonstrates that there is a difference between asking about how an object acquire its dispositions (here fragility) and what makes it fragile, something that depends on the properties of its structure.

The explanandum is, according to what is stated above the thing that should be in focus when determining what kind of explanation will best answer the question, so a scientist should always be careful and precise about what the explanandum of his theory is, and use the right sort of explanans accordingly. So the relevance of an explanation depends on what kind/part of a phenomenon the scientist wants to explain and can be entailed by the way the scientific question is posed. In many cases both types of explanations will be relevant for the sake of understanding the whole behavior of a mechanism or system, but depending on what kind of question is asked, again only one type of explanation might as well be sufficient to account for the problem under investigation. A constitutive explanation can for example be a necessary supplement to an etiological explanation by providing focus to the "wholeness" of the system by asking and explaining how a system has the capacity for a specific kind of behavior (Ylikoski 2013), or itself be sufficient when the phenomenon to be explained asks for how a system has the properties it has. This is often the case in neuroscience where the individual behavior of an entity or individual is not generally the focus of the investigation. In neuroscience, including the part that is interested in, and investigates consciousness, the point is often to gain a general understanding of how a structure can be responsible for and give rise to the phenomenon under investigation. In relation to what kind of explanation would yield a better understanding of consciousness, a constitutive explanation might seem better equipped to do the job. It could provide an explanation of what it takes of a system (the brain) to constitute, or realize conscious content. As this is important for how the Hard Problem of consciousness might one day be solved. I will return to this statement later, after discussing how a constitutive model is built up, and also how this is important for the possibility of explaining consciousness.

As stated above constitutive explanations give an understanding of how the parts in a system are constituted by explaining how the interaction and organization of the components and their activities relate to the system as a whole (Craver 2007). The constitutive explanation works by explaining the structure and organization of the causal components, and how these gives rise to a given phenomenon, the relation between the causal parts and

the constitutive explanation is thus asymmetrical. This means that the causal capacities themselves does not constitute its parts and their structure and organization. In other words the constitutive explanation explains the what in the organization and structure of the causal properties is like in the case of a given phenomenon, and the causal components and the relation between them does not themselves give a constitutive explanation. This means that the explanation works in a one-way direction, with constitutive explanations making sense of the causal relationship between the components. The constitutive explanation also describes a synchronous relationship, if there is a change in the basis of the structure that at the same time alters the causal capacities in the components there is no time laps between the two, they are not independent of each other and a change in one will lead to a change in the other, making "cause" and "effect" mutually dependent of each other (Ylikoski 2013).

The relation between the components can be seen as the entities in a mechanism, and the way they function (and thus the functioning of a given mechanism) can be explained by a constitutive explanation. As such the constitutive explanation can be understood as a relation of dependence, seeing as the causal capacities of the whole depends on the structure, organization and activity of the parts, both the organization and the activity of the parts do in other words have explanatory relevance for the constitutive explanation. Here see how the constitutive explanation builds upon the causal capacities found in the parts, and how the system has it's causal capacities thanks to the causal capacities of the parts. For short, the causal capacities make up and "build" the structure as a whole, and the role of the constitutive explanation is to explain the components, organization and activities that make up the structure of the complete mechanism. The adequacy of the constitutive explanation stems from its ability to explain the organization of the parts and how this leads to the causal activities that make up the whole of the system. But in the case of constitutive explanations, as in the causal ones, difficulties about determining the explanatory relevance of a given part in the system are likely to occur . This is especially apparent in cases where only some aspects of a structure are taken as a necessary explanandum of a phenomenon. In such cases problems of determining which of the parts relevant factors for the realization of the problem are should be accounted for. The fact that the structure of the mechanism under investigation might be realized by several causal capacities creates a redundancy of explanatory relevant factors for the constitutive explanation, but according to Ylikoski (2013) this problem can be treated much like the

problem of alternative causes in the case of causal explanations. In such cases where more than one possible cause can lead to the same effect what is to be held as the actual cause is not always readily available. When undertaking, for example, a thought experiment about possible causes of a given effect or event, there might be some problems linked to the determination of what is the actual cause in cases where several causes can yield the same effect. But in (natural) science, this is a question of empirical research, and the answer to what cause is actually the right one is often or always available through empirical experiments, and by looking more closely for the relevant explanatory factors. In fact there are close similarities between the problems found in causal and constitutive explanations, and these problems can often be solved in similar ways as well. So both the problems, the solutions and ultimately also the things they want to explain can all be similar for the causal and the constitutive explanations, even though the explanatory relevance depends on different things.

The question about how living beings comes to be conscious is posed best as an inquiry about what enables conscious experience. The Hard Problem of conscious- ness thus demands a constitutive explanation, en explanation that can provide an understanding of the underlying structure of consciousness and that can give a better insight into what kind of properties that are necessary for conscious experience to occur.

For several reasons, then, it does not seem to make sense to ask for a causal link to account for the relationship between the physical structures of the brain and the experiential contents of the mind. More likely the link, and so the best explanation is to be found by help of a constitutive model, aiming to explain and make out what the physical structure of the brain (along with relevant background states) must look like in order to give rise to qualitative consciousness. If this is right it might imply that the right kind of explanation for consciousness is constitutive. This notion is further supported by the thing that seems to be the core and the cause of the Hard Problem in philosophy of mind, the apparent difference in substance between physical structures and neuronal signals and the way this is experienced in our subjective consciousness. The metaphysical difficulty of making out and proving "logically" the necessity of a causal link between cause and effect is thus amplified by the substantial difference between the physical and subjective conscious entities. In normal science the remedy for the causal worry is empirical research and experimentation, but in the field of consciousness this is more difficult due to the many and complicated structures of the brain that need to be better understood themselves in order to give a better

understanding of their causal relevance. And this of course is topped by the seemingly insurmountable substantial differences between physicality and mind.

If the phenomenon under investigation turns out to not entail a causal relation- ship, a causal explanation of this relation might not be the best way to gain a better understanding of the problem. If it turns out that the relationship between the physical and the mental is better explained by a constitutive type of explanation, as is implied by this chapter, it would seem that the kind of explanation one should seek to close the explanatory gap of the hard problem is constitutive. The solution to the problem of consciousness should thus try to find out what kind of properties, structure and organization is necessary on the physical level in order for consciousness to "happen", rather than finding a temporal cause to effect solution.

3 Constitution and Consciousness

As discussed in the last chapter, the aim of constitutive explanations is to offer an understanding of a given phenomenon by appealing to the organization and structure of the parts that constitute a phenomenon or an event. More accurately, constitutive explanations are a description of the processes that underlies and gives the premisses for the occurrence of the given phenomenon. So instead of a causal explanation of why the phenomenon occurs we get a description that tells us what kind of properties are present, and how this is constitutional for the phenomenon or event. As such the constitutive explanation offers an understanding of the underlying structure of the phenomenon of interest (Craver 2007).

In philosophy and science of mind this type of explanation is seen as the way to go when it comes to finding an adequate explanation of consciousness. This is evident in the way cognitive science and philosophy of mind formulate theories of consciousness, where research first and foremost deals with, and tries to find, a constitutive explanation of consciousness (Craver 2007). As we have seen it is difficult to expect, or even ask for causal links between physical indications of mind, and the felt experience of the conscious content, due to the apparent difference between physical substance and the abstractness[4] of conscious experience. Examples of this kind of constitutive approach in the field of philosophy of mind is found in many theories such as representational theories, functionalism (Putnam 1967, Fodor 1968), identity theory (Feigl 1958, Smart 1959) and so on. Although theories in areas such as these do not always make explicit claims of providing constitutive explanations, I will try to make it clear in what follows the kind of explanations found here are profoundly constitutive in their form. But as the mystery of consciousness has not yet been solved, large differences can be found between the different theories of how conscious- ness might arise, and accordingly what should be seen as constitutional factors for this.

The mind/brain identity theory for example, is a form of materialism or physicalistic theory of mind.[5] As the name of the theory indicates, the mind/body identity theory says that a state of mind is identical to a brain state, i.e. mind and brain correspond to each other, and for all mental states there are necessarily some brain state that can be

[4] By abstract, or rather the phenomenal nature of the mind, I refer to the difficulty of making objective measurements of conscious experience.

[5] With a physicalistic theory of mind I refer to theories that support the notion that mental states can be reduced to physical states in the brain.

identified with it. The phenomenal experience of the world, our thoughts and feelings, can thus be explained by appealing to physical brain states. To describe the mental state of interest, a supporter of the physical mind/brain identity theory would turn to the states found in the brain in order to give the appropriate answer as the mental state is identified/identical with this physical state. In the physical mind/brain identity theory a constitutive explanation of consciousness would thus be focused on explaining the properties of the brain state that identifies the mental state. A constitutive explanation will in this case depend upon a specific explanation of the brain and brain states, since this is what constitutes mental states. Here identity as it is found in the mind/brain identity thesis is understood as a special case of constitution, where a constitutive explanation of mental states depends on an accurate description of the brain states identical to the mental states.

Physicalism, the branch of philosophy of which the kind of physical identity theory discussed above is a part equates mental states (mental content) with a physical state or process, so that there is nothing to be found out about the mental states over and above what can be known about this state in physical terms. In this case the two types of states are identical to each other. There is in other words a strong correlation between mind and body, to the point where, according to Armstrong (1968) all mental states can be seen as corresponding to a physical state which should be perceived as literally identical to it. According to Place (1945), one of the founders of brain/mind identity theory, "The logical objections which might be raised to the statement 'consciousness is a process in the brain' are no greater than the logical objections which might be raised to the statement 'lightning is a motion of electric charges" (Place 1945, p. 255). The statement might benefit from some moderation[6], but it still indicates in a good way how supporters of the mind/brain identity theory logically conceive of consciousness and the relationship between consciousness and brain states.

The identity theory is to some extent based upon the theory of neural correlates of consciousness. But instead of looking for some causal relation that binds the two together, or seeing consciousness as an epiphenomenal byproduct of neural activity as is the case with Cartesian dualism and epiphenomenalism, the Identity theory takes mental activity to be the same as activity in the brain. To find out something about consciousness one will in other words have to study processes in the brain, as the expression of consciousness can be

[6] Feigl (1958) and Smart (1959) later somewhat moderated Place's statement, indicating that lightning and electric charge, or sensation and brain state for that matter might not have the same meaning, yet they still refer to the same thing.

found here. The theory of mind/brain identity is a good one, as it gives a concise answer to what consciousness is in a simple way without treating qualia as something intrinsic and abstract. Instead qualia is treated the same way as every other mental state, as a physical state in the brain. But this view has proven to be problematic and has been criticized for being too rash by amongst others Chalmers (1996) in his famous zombie thought experiment. With the zombie argument Chalmers discusses the metaphysical conceivability of a person that have all the same physical states as a normal, phenomenally conscious being, yet the person lacks any experience of these physical states, and so is without consciousness. If two people could have all the same physical states, as well as the same behavioral reactions to sense input, with only one of them having any conscious experience of this, it implies that there is something more to consciousness or qualia than the physical states of the mind. This argument have been disputed, as it is a thought experiment, and not empirically verifiable. An opponent might argue that it is actually not possible to imagine such a person, and that if the brain states of the "zombie" are the same as the normal persons, then this implies that there also are corresponding mental states, and this person can in fact not be an unconscious zombie.

3.1 The Problem of Multiple Realizability

When mental states are understood to be identical with brain states this also makes the way for the advantage of dealing with the problem of dualism in a straight- forward way, as it skips the whole problem by identifying consciousness with specific states in the brain. When there is no contradiction between mind and brain there is no room for dualism, as the two, the mental and the physical, are just different expressions of the same thing. From this it is easy to see how the mind/brain identity theory is a theory under the physicalism parole, as mental states can be explained by the states found in the brain. But the theory of mind/brain identity and the position of taking consciousness as a purely physical instance in the way it is done here is somewhat problematic according to some philosophers. One of the main objections to the theory was presented by Hilary Putnam (1967), he stressed the fact that if mind and brain are to be understood as the same thing, and that the brain state (and thus the mental state) is characterized by being physically unique, only creatures in this specific physical state will have the corresponding mental state. This will exclude all creatures that

are not in the given physical state from having any mental states. This definition might prove to be too narrow, as it is fully possible to imagine that mental states can occur without such a specific physiological and chemical structure. In fact it is not really certain that all human beings are in the exact same physical-chemical stats when they experience a sensation as for example pain, so the identity theory that states that there is a highly specific physical state for every mental state might be too forceful. This has been thoroughly discussed in William Bechtel and Jennifer Mundale in their 1997 paper "Multiple Realizability Revisited".

The critique of the mind/brain identity theory as it is put forth here by Putnam can also be used as an argument for functionalism and the functional understanding of how mind (and maybe also qualia) is constituted. One of the arguments for functionalism are based just on the hunch Putnam describes here, that there might be no "equal" brain states, and thereby no equal mental content. Instead of identifying mental states with physical brain states, he proposes that functional states, as an expression of the causal relationship between sensory input, other mental states and behavioral output are what identifies and constitutes a specific mental state.

An argument as to why the explanation Putnam gives for his dissatisfaction with the physicalism of the identity theory might be to narrow is according to Bechtel and Mundale (1997) the way philosophers of mind tend to describe brain states. According to their 1997 article "Multiple Realizability Revisited" the normal way of describing a brain state differ substantially from how the same phenomenon is described by cognitive scientists. The difference lies in the definition, with philosophers (at least in many cases) referring to a brain state as the chemical-physical state of the brain at a given time, in the same way Putnam does in the example noted above. Scientists who work on the brain has on the other hand traditionally not been concerned about detecting and identifying "brain states" (as these are extremely fine grained, and difficult to detect or even possible to use as grounds for relevant information (Bechtel & Mundale 1997)) the task of the scientist has rather been to find out which area of the brain can be connected to certain mental states.

This kind of neuro-anatomical brain mapping, has been performed by scientists at least since the 19th century, and the project has largely been focused upon finding which areas of the brain are associated with given mental states such as for example hearing and memory, with a special focus on the study of vision. In this way cognitive scientists are in good hopes of finding neural (or at least area specific) correlates of consciousness. One of the arguments of Bechtel and Mundale is thus that it is possible to find some sort of identity between

mental states and brain states, here in term of activity in a certain area of the brain. As research in this area progress one can imagine the mapping of the brain to getting more and more specific in relation to mental states, making a finely grained discrimination of a brain state more likely to be found.

This relates closely to what Bechtel and Mundale see as the primary error done by philosophers in relation to multiple realizations, and the reason for why it is held by many as a viable possibility. The two scientists insists that the possibility of multiple realization depends on philosophers tending to have an imbalance in terms of how finely grained mental states and brain states are taken to be in relation to each other. Often brain states are taken as more finely grained than the mental states, meaning that philosophers are looking for very small and demarcated brain states to cover mental states that are actually more general. In other words they might imagine and look for very small and finely grained brain states to be necessary in order to identify wide mental states such as hunger or pain. This imbalance is an error in terms of how the brain functions empirically, and to get a clearer picture of how brain and mind relates to each other physical and mental states should be kept as having the same "grain size" when they are compared with each other.

Bechtel and Mundale stress the importance of keeping the size of the brain states on a par with the mental states because this lends support to their main argument, that multiple realizability if it turns out that the same areas of the brain are responsible for the same mental states between species (Bechtel &Mundale 1997). The point is that cognitive research on brain mapping shows that some areas of the brain can be linked to certain mental states, and that these correlations hold for several species. The same brain area can in other words be recognized across both individuals and different species. As the mental states become more finely grained they will again relate to smaller areas of the brain, making more finely grained mental states dependent on more specific parts of the brain all the way down to chemical and molecular structures. A typical argument from the side of the supporters of multiple realizability is that there can be no specific brain states that correspond to a given mental state due to lack of functional evidence that links the brain state to the mental state. But if this is due to the imbalance of the grain size used to describe the two this might serve as a reason for why multiple realizability on the physical level is impossible as long as the neuroscientists can identify at least coarse grained brain and mental states with each other. According to Bechtel and Mundale (1997), the only reason why it is sometimes difficult to find the appropriate relation between a mental state and a

brain state is that the division one is operating in is to coarsely grained, and all that is needed to find the desired relation is to make both the mental and the physical state specific enough. This might seem difficult at the present time, but if neuroscience and work on brain mapping keep evolving it might be conceivable that even the smallest differences in neural states can be detected. If no such corresponding brain states (that correlates to given mental states) can be found between species, the reason probably is, if one follows Bechtel and Mundale, that there really are no corresponding mental states between the two. Two clearly different species such as a cow and a human being can probably share the mental state of being hungry, and as such one would expect to find some similarity between the brain states of the cow and the human being, for example in the form of neural activity in the same area of the brain. But at a more finely grained level there is really no knowing whether brain states can be common for different species, and if not, the individuals of the different species will not share the same mental states either. This is no surprise, and should not be seen as a problem according to Bechtel and Mundale, as different species will probably have different dispositions for how to perceive and behave in the world, and if there is no equality between the brain states of different species, there is probably no point in expecting to find equal mental states.

The possibility of multiple realizations should thus not be expected from neuro-scientific evidence, even though it seems intuitively plausible from the functionalist point of view. The error committed by the philosopher of mind is to not adequately differentiate between mental states and brain states, taking it for granted that fine grained brain states can correspond to the more covering mental states. In other words psychological functional states are not multiple realizable on the physical level according to modern neuroscience, a scientific branch that remain hopeful of finding neural correlates of consciousness. It is on the other hand not in Bechtel and Mundales (1997) to disprove functionalism as an overall theory, they merely wish to point out the difficulty that should be associated with multiple realization between the physical and the mental level. From the functional level there is still, on the other hand, room for multiple realizability of mental states, and as such the functional states should be seen as important pointers to the kind of mental state that is apparent in the subject under research in order for a corresponding brain state/area to be detected.

Early neurobiological research on brain mapping was done by experiments and studies of animal brains, which later was generalized to also account for the human brain (Bechtel

and Mundale 1997). This kind of generalization might seem problematic, but scientific research on the area shows that basic neural mechanisms seem are similar in all animals, and such studies has indeed proven to give a better understanding of the human brain (Sylwester 1995). Research on the human brain and brain mapping has also been done with lesion studies on human beings who have acquired brain damage, and in more recent times brain mapping on humans has been studied with brain scanning such as PET, MRI and fMRI. With these methods a map of the human brain able to show activity in somewhat specific and separated pats of the brain is possible to make, and this can again be used as evidence for a correlation between physical and mental states.

The thing that explains consciousness according to the mind/brain identity theory is the physical and chemical structure of the brain, so what constitutes conscious- ness here is the physique of the brain. But as Bechtel and Mundale (1997) implies in their article, even if multiple realizability is wrong, (possibly making mind/brain identity theory seem like the better option when it comes to plausible explanation of mental states) it seems to be no real hinder for the notion that other theories such as functionalism are significant for how to explain mind.

3.2 Functionalism

As mentioned above, functionalism[7], unlike the physicalistic identity theory, does not take the physical or chemical structure of the brain as the basic constitution of consciousness, rather the identity, or constitution of consciousness should be looked for in the way mental states function. Instead of looking at what kind of neuronal structure or brain state corresponds to a given mental state, the functionalist looks at what kind of role the mental state plays in the system of which it is a part. The function of a mental state can be for example to utter sounds and in other ways show signs of discomfort when a subject is in pain. An illustration of this can be made by a comparison with a biological concept, for example a kidney, where the function of a kidney is determined by, amongst others, its ability to filtrate blood, and not by its physical features. The same goes for the notion mental states,

[7] I later distinguish between what McLaughlin has called Role-functionalism and Role filler-functionalism (McLaughlin 2006), and the way functionalism is described right here can possibly be said to only apply to Role-functionalism.

it is determined by its function rather than physical features in the brain, when following the functionalist view on consciousness and mental states.

Still there are two separate ways mental states can be identified within functionalist theories. A separation goes between role-functionalism and what McLaughlin (2006) called filler or role filler-functionalism. When deciding how to identify a mental state, some functionalist thinkers refer to the higher level property of the causal relation between input, output and other mental states. McLaughlin calls this role-functionalism, because the functional state can be understood as playing the role of the mental state. Here the mental state is identical to the functional state itself. The other way for functionalism to describe the nature of mental states is by viewing the functional property as a descriptive term, a definite description for the lower level property realize these functional states. Here the lower-level or physical property that gives rise to the functional state is seen as playing the role of, or realizing the functional state. The theory of role filler functionalism is thus a reductionist theory, and have close links to physicalistic identity theory, but differs by using functional states as an intermediary factor in the description of mental states. In accordance with the Bechtel and Mundale (1997) theory, the physically reductive role filler-functionalism should not be seen as multiple realizable, but if the reference to physical states is avoided in the description of mental states, as is the case with role functionalism, multiple realizability is still a possibility.

Functionalism derives, and lends some support from several philosophical and psychological branches, such as behaviorism and analytic philosophy, and was partly constructed as an argument against the rigid rules for the constitution of mental states found in the identity theory, by one of the functionalist pioneers H. Putnam. As mentioned above Putnam understood the identity theory as unable to suffice as a plausible theory of mind, as it did not allow for mental states in anything other than physically (and chemically) identical structures possibly only the human brain. As this position does not seem to cover all the ways we can imagine mental states to manifest themselves, Putnam (1967) understood the theory as giving evidence for the multiple realizability of mental states. A different theory for identifying and describing the nature of consciousness was thus needed. In role functionalism multiple realization is a possibility due to the fact that consciousness or mental states are equal to their functions, the identification of a given mental state depends solely upon its individual relation to other mental states, along with sense input and behavioral output. In other words the physical structure in which these relations form is not

essential to what the mental state is like. One might say that functionalism is a computational model of consciousness where only the software is constitutional for the mental states, and the hardware is optional as long as its structure are able to provide the right functions. A brain made out of silicone might function just as well as a brain maid of human cells,

As long as it provides the same sort of functions. As the example with the kidney illustrates, it just does not matter what kind of matter an organ (or other entity) is made of, as long as it is sufficient for the thing to fulfill its task, or rather function in a given way. In role filler functionalism on the other hand, there is room for the physical states of the brain, as these are seen as the realizers of the functional states. As long as it is a point for functionalism to provide an explanation for the occurrence of mental states, where there is room for the possibility of somewhat different functional states giving rise to more or less the same mental experience, and if Bechtel and Mundales (1997) argument against the possibility of multiple realizability from a physical level is correct, the role functionalism might provide a better option for how functionalism should be understood.

In addition to these two ways of understanding the role of functional states and their relation to mental states, functionalism has several different varieties, where the functions of mental states are defined in different ways. According to amongst others Putnam (1967), the mind can be understood roughly as a machine which uses input and the dispositions found in the machinery to produce output. This output might translate into the function of the machine, in the same way as mental states can be seen as functions of the causal relations mentioned above. This model loosely resembles the psychological theory of behaviorism, as it focuses mainly on the output of the "machine"/brain as the function of the mental state. Another sort of theory is found in psycho-functionalism (Fodor 1968), which takes empirically available data about human psychology, often cognitive psychology as a basis for how mental states "works". Psycho-functionalism use the methodology of the psychological theory to determine the function of the mental states, and by using framework, it makes the scientific, cognitive structure fundamental for the function of the mental states.

Other variations of functionalism also exist[8], but for my purpose here it is enough to make out that within a functionalistic framework there are several ways to understand how the functional states of the mental states should best be described. The constitutive

[8] See for example Kim (1989), Fodor (1968) and Kripke (1980)

explanation of mental states will thus vary between the different theories. Most of the functionalist theories does however have some common problems, amongst others the theories are often constructed in such a way that they exclude individuals who lack certain functions from having the corresponding mental state. Something, or someone who for example does not have the given functional expression of for example the mental state of being in pain, is seen as to lack this mental state on a whole (Block 1980), since a functional description is based upon functional expressions, it is not able to provide an explanation for mental states where such expressions are lacking. This is damaging to our notion that we can understand and share mental states, and it is thus a problem for functionalism because the theory in this case does not give a sufficient explanation of how mental states can be understood and identified.

More importantly (for the purpose here), and more damaging to functionalism, is the problem of qualia. This problem has proven to be especially difficult, and even though functionalism has more problems (as seen above), it seems like none of these has been as difficult to refute as the problem of qualia. There are here several worries, and Chalmers' problem of absent qualia has already been presented. Another serious critique of functionalism takes up the problem of inverted qualia, and is largely connected to work done by Ned Block (1980). The notion of inverted qualia express the worry that it might be possible for some individuals to have an experience of something that is opposite to how other people experience the same thing, or maybe even opposite to how the person itself have experienced the same thing at an earlier time. The most popular way of making this argument is by imagining an inverted color spectrum. The inverted spectrum thought experiment is the hypothesis or the thought experiment that you can imagine other people to see colors differently than you do, for example we could perceive of a person where every color has been swapped out by its complementary color. This means that the person in question sees yellow as blue, green as red and so on. The notion is not a new one, and many people, even children often wonder whether other people see colors the same way they do. The fact is that we in many cases actually don't! More often than not color will appear in a more or less different shade of the same color from person to person, with the difference growing with quantity of different (biological) factors (Block 2006).

In his argument Block imagines that the wires in the visual system has been crossed, so that the connection from a cone that detects red light are transferred to a part of the brain that gives the mental image of green, either all the way from infancy, or as the

consequence of an operation performed later in life. In this case the person with the "crossed wires" would perceive differently of colors in relation to people with a "normal" color perception. But how could anyone know this to be a fact? In the case were the person with the inverted spectrum has undergone an operation the person could probably give some sort of expression of perceiving as red today the things which appeared green yesterday. The person could point to the grass and tell us that it now looks to have the color of a ripe strawberry, at the same time as knowing that this was not the case yesterday, when it was green (we can't actually know what green was like to this person earlier, maybe it is first now that the colors actually look "normal", the way they look to me?). A person that, on the other hand is borne with an inverted spectrum would probably not be so inclined as to see that there is anything wrong or abnormal with the way he perceives the color of grass. He or she would probably refer to the grass as having the color green, even if it would look red to us if we could take a look inside his head. Block stresses the point that in such a case nobody really has the right to say that the way he or she experience color is the right way, as there is no clear reference or standard of which this can be measured. If the way other people experience a color is hidden to us due to the impossibility of communicating the difference between the two, there is really no way of knowing who is right about what the actual color of a thing looks like.[9] . Additionally we might imagine an experiment where a person putts on a pair of goggles that inverts the color spectrum so that every color looks like its opposite (complimentary color) when they are put on. In a similar experiment done with glasses that makes the world seem like it is upside down, George M. Stratton (1896) found that after wearing the glasses continuously for several days his vision gradually adapted, and turned the image over so that the world again appeared the right way up. When he then removed the "upside down" goggles, the world again seemed to be turned upside down, even though it now actually was turned the right way, making him disoriented once again. In this case the brain of the person undergoing the experiment (Stratton), adapted to seeing the world upside down by reconstructing the image to look like it normally would, and as in the case of the inverted spectrum rewiring, after adaption both the functional and the mental states are the same as before the goggles were put on, even though the image that actually hits the eyes has been reversed.

[9]Here the look of the color refers to the way it is experienced of a given individual. If color is instead described in physical terms as light with different wavelengths, a color is determinable in reference to its specific wavelength. What kind experience this yields in different individuals is a different question, and is what I discussed here.

The hypothesis of inverted spectrum is an argument against functionalism. The key here lies in Block's attempt to show that it is impossible to discover the functional difference between an inverted and a normal color spectrum. If it is possible to invert the color spectrum in a human being, so that its reversed in relation to another without there being any detectible functional difference between the states, functionalism cannot explain the difference between the two states.

This problem shows up due to the fact that the purely relational character of the functional states makes it impossible to determine what kind of experience an individual actually has in connection to his mental states, as long as the functional expression is the same, all one can be certain about is that the mental state is the same. In other words it is here possible to imagine someone having an inverted color spectrum, where the functional expressions correlates with the opposite colors of what is "normal", making the actual experience or qualia an unavailable factor for the functionalist theory. This also opens up for the notion of absent qualia, supporting the notion that there might exist individuals without experiential content, who actually lack the qualitative portion of mental states. This notion points to a deep problem for functionalism and its relation to qualia, because even though functionalism is happy to explain mental states in terms of the functional states, these states are characterized by the extrinsic relationship between perceptive input, output and other (causally relevant) mental states without making an appeal to the felt quality of the intrinsic these mental states. In other words functionalism are only in the business of explaining the extrinsic properties of mental states, and shows little or no interest in the intrinsic properties of qualia. A reaction to this was proposed by John Haugeland (1978), who suggested that the theory of functionalism can be narrowed down to only serve as an explanation for mental states that lacks qualitative content, making qualia absent from the functionalistic explanation. This again relates to the fact that even though every functional state is equal to those found in a "normal" human being, there is still no guarantee that there is a qualitative experience connected to it. Ned Block (1980) illustrates this by imagining people substituting as the neural structure of the brain, duplicating its functional organization by messaging sensory inputs from an artificial body, and sending messages between each other much like a normal brain would do. This could yield functional states equal to those found in the normal brain, but it seems that it would not give rise to any form of experienced qualia over and above the qualitative experience of the individuals participating in the experiment. This moment of critique will be

investigated further in the next chapter Here I argue that there might be a way for functional states to include this qualitative aspect, if we supplement functionalism with a phenomenological understanding of how consciousness is constituted.

These arguments are damaging to the functionalism in as much as one takes qualia to be a real thing. This seems to be the attitude of most scientists and philosophers of mind, as it is natural to take the way we "feel" and experience our conscious content to be real. Still there are those who are inclined to throw away these kinds of felt qualities and experiences all together in order to avoid problems like those mentioned above (along with several others), mainly eliminitavist theories of consciousness (see Dennett 1978, 1988 and Churchland 1981). According to this kind of theory people is in a way understood to be "zombie like", as there is actually no experiential content associated with the neural activity of the brain.

The "zombie-argument" (Chalmers 1996), that previously was used as an argument against identity theory is also used as an argument against role filler functionalism. The problem with the role filler type of functionalism is that it identifies functional states with physical states, and physicalism gives rise to the possibility of the existence of philosophical zombies of the type mentioned in the eliminitavist theory, human beings who live and have physical and functional states without qualitative and experiential counterparts. The zombie possibility seems highly unsatisfactory if we want to hang on to the notion that human beings are generally conscious and able to experience the contents of our mental lives (For details see Chalmers 2002). Levine (2001) later proposed that the conceivability of such zombies might not make functionalism and physicalism false theories, as physical and functional states still might serve as necessary constitution for consciousness. In relation to qualia however the theories seem to open up for the explanatory gap, as they do not sufficiently explain how the qualitative and experiential factors of consciousness arise from such physical and functional states.

The Hard Problem can in other words not be solved by describing mental states as purely physical or functional states. Both the mind/brain identity theory and role filler-functionalism can be described as reductive theories, because consciousness is perceived as something that can be reduced to physical states in the brain. Role-functionalism is on the other hand not a reductive theory in this sense, as it does not identify mental states with the states found in the brain. So there is in other words a distinct difference between Role filler and role-functionalism, as one of them reduce mental states to physical states in the brain and

takes the functional states as a denoting factor of this relation, while the other actually identify mental states with these functional states. But it might also be possible to take role-functionalism as a sort of reductive theory. This is because it holds that mental states can be described in terms of, or as, something else, namely the functional states. In eliminitavist theories on the other hand, the existence of consciousness is denied on a whole, and if this notion seems to be unsustainable (as it does to many) it is necessary to look a bit further in order to find a viable explanation to consciousness. This is the goal of the physicalistic mind/brain identity theory, but this theory also seem to be insufficient as it does not provide a necessary explanation for the qualitative aspect of mental states, as having a certain physical state does not entail having conscious experience. The notion of qualia are more or less excluded from the discussion in the physicalistic identity theory, neither is it logically implied. Functionalism ends up in much the same way, as it does not offer an explanation for the intrinsic character of qualitative experience. In the case of inverted spectrum it is not even clear that the theory offers a sufficient explanation to any mental states. This implies that there is still some way to go before the occurrence of the qualitative aspect of mental states can be thoroughly explained and understood. Further discussion is needed about what constitutes a mental state in order to come closer to a conclusion. If, as it seems here there are many ways to understand the constitution of consciousness, and disagreement between different theories about what the constitution of consciousness is, and how it is constituted, a better understanding of consciousness might be needed. In what follows I will take a look at how a phenomenological perspective on consciousness might open up for a deeper understanding of what consciousness really is, along with an alternative view on how consciousness can be explained in constitutional terms.

4 A Phenomenological Understanding of Consciousness

In order to give a constitutive explanation of consciousness, it is necessary to have a full understanding of the phenomenon of the explanation, namely consciousness itself. In this paper I have referred to consciousness as what Nagel (1974) called qualia, the inert, subjective experience following mental states. But an even clearer description of how consciousness should be understood is to be found in Husserlian phenomenology. The theory he proposes is also ready to deal with the relationship between consciousness and the outside world, something that has been one of the main issues in the philosophy of both Nagel and Chalmers. The mission of Husserl's phenomenology is to analyze the structure and meaning of human psychological concepts in order to get a clearer understanding of the nature of consciousness, and how we are able to perceive the world. The project aims at disclosing what he describes as "pure consciousness", or what is left when one abstracts away from, and excludes (thoughts about) the world and self. The point is to perform a phenomenological reduction, where the world is bracketed[10] in order for conscious awareness to appear on its own (Husserl 2012). As such, consciousness and the way it is directed at the world is in itself open for research in a different manner than what is the case in traditional analytic philosophy and cognitive research. This is because the phenomenological reduction lets consciousness, instead of the physical word be the phenomenon of research and interest, as the world "in itself" is taken out of the discussion.

In the Husserlian way of investigating consciousness, the conscious experience itself is the starting point for the evaluation, and the way of doing research is more or less turned around in relation to the method of analytic philosophy and classical neuroscience. In a way the tables have been turned, so where conventional philosophy of mind talks about what a plausible constitution for consciousness might be, phenomenology has the habit of asking what the constitution of the conscious mind might be like in order for us to experience the world as we do.

The phenomenological way of perceiving consciousness and its relation to the world has also been adopted and developed by other thinkers, mostly belonging to the continental

[10] The Phenomenological reduction: By performing an epoche, the world is "bracketed" so that the focus can stay on how the world is constituted in our minds.

tradition, such as Heidegger (2007), Sartre (1957) and Merleau-Ponty (2012). In recent times the phenomenological tradition has been further developed by philosophers such as Shaun Gallagher and Dan Zahavi (2008), and Evan Thompson, Luiz Pessoa and Alva Nöe (1999), with the mission of giving a better understanding of how consciousness should be perceived in philosophy of mind.

In traditional analytic philosophy of mind we meet several different theories of how consciousness is constituted. Different explanations of consciousness are thus available, but as we saw in the last chapter, all of these are still open for critique, and there are some profound problems associated with more or less all of them if we want to explain the Hard Problem, while still using description of consciousness as it is stated by Nagel (1974) and Chalmers (1995). The main objection to both physicalistic (reductionist) theories and, to some extent functionalism, is that they do not provide adequate evidence for, or even an understanding of the experiential or qualitative aspect of mental states. In phenomenology on the other hand, the conscious or experiential part of mental states is where any theory seeking to explain consciousness should start out, and in order to give an adequate explanation it is necessary to first understand what consciousness is. Because of this difference in terminology and the way consciousness is understood, it is interesting to look at how the two methods or traditions fit together, and whether one can be used to supplement the other in order to gain a more coherent and wider reaching explanation of consciousness.

In this chapter I am going to take a closer look at the phenomenologically inspired enacted theory (fronted especially by Alva Nöe (2009)), and how this somewhat opposite view on what is explanatory relevant in the case of consciousness might be used to supplement analytic philosophy of mind. Here I will mostly concentrate on functionalism as it is explained in the previous chapter, and look at whether the phenomenological understanding of consciousness can be used to supplement this functionalistic effort to answer the Hard Problem. As discussed in the earlier, functionalism is unable to account for the experiential aspect of consciousness because the functional states that identify the mental states do not express these intrinsic qualities. But with a supplement given by the phenomenological description of consciousness, it is possible to describe the functional states as also expressing the subjective experience of consciousness. This "new" description of mental states can hopefully, in turn, be used to supplement functionalism by providing an objective or extrinsic referent to experiential mental states. As proposed by Haugeland

(1978), there might be a job for someone else than functionalism to explain the experienced quality of mental states, and if phenomenology can provide any kind of insight on this area, the two theories might supplement each other nicely, and as such, give a plausible explanation of qualia and the mind/body problem. An example of this is found in the work of, amongst others, the Italian neuro-scientist Giulio Tononi (2000). In his information integration theory Tononi proposes that a science of consciousness must be based on a fundamental understanding of what consciousness is, and what it does. The insight into what consciousness is like is gained through phenomenological investigation of the intentional character of the mind in relation to the world. The project is aimed at providing a system for the understanding of phenomenal consciousness through mathematization, which in turn enables neuroscientists to identify necessary components in the brain. The project bear some methodological similarities to the earlier developments of the Canberra plan, based on the work by amongst others Lewis (1970), Armstrong (1968) and Jackson (1998). The similarity between the two is the effort to ground a neuro-scientific explanation of the mind on an intuitive understanding of consciousness, similar to that of phenomenology. The project aims at disclosing the a priori structure of thought we have when we engage with the world. The point of the analysis is to map out what is required in order for experiential consciousness to be realized.

The philosophical direction taken by the Canberra Plan and the thinkers who support it is that the first task of a philosophical project is to provide this kind of a priori analysis, rather than to make up synthetic theories about how the (physical) world should be apprehended. In this way the theory is focused on finding/analyzing the a priori categories people use in their everyday thought of the world and mind. The point is to undertake such an analysis in order to make clear the kind of structures that are applied when we (conscious beings) relate to the world, and this again can be helpful in terms of making out how the world is categorized in "normal", folk-psychological thinking. Further, the kind of insight given through the a priori analysis should lead to a better foundation for making synthetic theories about the constitution of the world (and mind). The project can be compared to, and as such give some support to the functionalist role-filler theory, because it points out the a priori categories of human thought. The understanding of these categories is used to structuralize and make sense of the world in order to determine how consciousness should be understood and explained in order for these categories to be realized. This is similar to what role filler-functionalism does when it identifies mental states for the sake of finding the

underlying and realizing physical states. Both theories make a point out of the necessity of being able to clearly describe the explanandum in order to make a thorough explanation of the phenomena of interest, and this is in line with the phenomenological way of accounting for consciousness.

Part of the reason why functionalism does not by itself provide a complete explanation of experiential consciousness was discussed in the previous chapter. An important point being that functionalism does not extend the discussion of mental states to include experience or qualia (Livingston 2005). This probably has something to do with the fact that role functionalism lets psychological states stand as constitutional for the appearance of consciousness, in order to stay on good terms with the natural sciences and empirical research. But to describe mental states in terms of the relationship between psychological states such as behavior and sensory input does not necessarily grasp the experiential aspect of consciousness. By appealing to the extrinsic qualities described by classical psychology, functionalism keeps within the framework of cognitive science and physicalism, and meets problems as these theories do when it comes to explaining the intrinsic character of qualia. As illustrated with objections stated by Ned Block (1980) in the last chapter, functionalism is able to explain (or give grounds for) mental states, by appealing to the relation between psychological states and behavior, but the problem is that neither psychological states nor behavior express subjective experience.

This is what is reflected in Blocks critique in the inverted spectrum hypothesis. So, even though functionalism provides a good framework for how mental states can be understood, by showing how these mental states can be identified with functional states, this is not enough to provide for an explanation of the experiential consciousness. In other words the theory fails to provide an adequate explanation of the Hard Problem. Does this leave an opening for phenomenology and the understanding of consciousness provided for here? There certainly seem to be some good evidence for thinking that functionalism and phenomenology have something in common. They both distance themselves from a pure physicalistic reduction of the mind[11], and they both appeal to logical analysis as a method for describing mental states. As such both are less close to empirical methods of scientific research than a functionalist might like to admit. But this is not necessarily a devastating problem for functionalism, as it depends on how the constitution of consciousness is understood.

[11] At least this is the case for both phenomenology and what I in the previous chapters referred to as role-functionalism. Role-filler functionalism is on the other hand more in line with a physical reduction of mental state.

One of the goals for many physicalistic reductive theories of mind is to find neural correlates of consciousness in order to show that mental states are equal to, or at least dependent upon physical states in the brain. This is an important notion in cognitive research on the brain, and by finding such a link researchers hope to find the connection between the mental and the physical, and thus an answer to the mind/body problem. This stands in strong contrast to the enacted theory of mind as it is presented by Alva Nöe (2009), who, contrary to the beliefs previously discussed in this thesis, urges the position of the interaction between body, mind and world as constitution and explanation of consciousness. As a base for the enacted theory, Nöe questions the foundation of modern neuroscience by pointing to how this foundation is nothing more than an assumption, an assumption that the base and constitution of conscious experience is a neuro-scientific phenomenon. According to Nöe there are many reasons, and good evidence for his attitude, the main reason being the fact that no direct link has been found. It is of course possible to find some correlation between physical and mental states, as is proved by amongst others brain scanning and lesion studies, but this kind of knowledge says nothing about how mental states actually are, or at least could be realized by the neurological state. As this conception of the relationship between mind and brain does not give an explanation for the phenomenon of interest, namely consciousness, the whole theory stated by physicalism is degraded to nothing more than a, in the eyes of Nöe, false hypothesis or assumption.

4.1 Nöe's Enacted Approach

The alternative proposed by Nöe is, that instead of thinking about conscious- ness as something purely neurological that somehow mysteriously arise and manifests itself in the brain, consciousness happens through interaction between several other parts than neurological ones. One of the reasons for this is that neuroscience itself seems to be quite far from finding a solution to the problem of consciousness, for as we have seen, even though such cognitive research can be used to recognize correlation between neural and mental states, the mystery has only been moved away from the brain as a whole and towards the neural structures inside the brain. Still the problem is the same, how can such

physical states possibly give rise to conscious experience. This might also be illustrated by appealing to what Ned Block (1994) referred to as a difference between access consciousness and experiential or phenomenal consciousness. The difference is that access consciousness is equal to those mental states that are accessible to functionalism, or the extrinsic qualities of mental states, while experiential consciousness is characterized by being intrinsic subjective states, the way things feel to us and the way we think about something. This difference speaks in favor of thinking that it is possible to talk about more than one type of mental state, where the first kind is responsible for the extrinsic expression described by functionalism, and the other for the intrinsic and felt quality of experience. The physicalistic hypothesis that all mental states somehow spring out of the brain is thus too rash, as it is unable to account for states of the second kind. It is this inability to explain the experiential aspect of the mental states, on the part of all the physiologically grounded theories, that makes Nöe propose that if one wants to find out what consciousness really is, it is necessary to look further than, and outside of the brain. The trick is, according to Nöe, that one should stop thinking about consciousness as something that happens inside the brain at a specific time and place. Instead consciousness should be perceived as an ongoing process, set up by the relationship between the mind, body and environment.

To demonstrate this notion he considers an example of the typical "brain in a vat" thought experiment. If classical neuroscience is right in the hypothesis that consciousness arises from the neural network of cells in the brain, a network of cells wired up in a suitable manner should be able to give rise to consciousness even without the context and the body of a living human being. In such a scenario there could, by the first look of it, in other words be conscious brains in a vat. But is this really all it takes? According to Nöe, in order for consciousness to occur the brain cells in the vat would probably need to resemble a real brain in order to have the right components and structure that make consciousness possible. Additionally, the vat that provides the surroundings of the brain would need to provide the right kind of nutrition and the like for the brain to work, and something would additionally need to provide some kind of sensible input for the sake of creating something to be conscious about. This scenario seems quite familiar as it clearly resembles how consciousness arises in a living body, and this, the presence of life, might be exactly what is needed for the constitution, and thus realization of consciousness, even in the brain in a vat scenario. Nöe's conclusion is that consciousness is made possible due to the fact that we are

alive, and as such it is not plausible that research on the neural system and brain cells alone will yield any significant insight into how consciousness is constituted.

So consciousness is something that is, at some level, constituted by life, and not necessarily only the brain. This is also evident form research on brain imaging, and the kind of information we can hope to get here. Even though it is possible that brain imaging such as fMRI and PET scans can show some correlations between the brain and mental states, there is no direct explanatory link between neural activity as it shows up in the scans and consciousness. At least if we are going to believe some of the empirical studies done in the field (A.M Owen et.al. 2006), on this link between consciousness and neural activity. They found, through numerous studies of patients in a so called vegetative state, where there is no evidence of consciousness to be found by neither scanning or personal reports, that the lack of consciousness at this one point in time does not entail that the patient is really brain dead. For surprisingly it is possible for such patients to regain and report full consciousness, without regaining the full kind of neural activity they possessed before going in to the vegetative state. Neural activity detected through brain scanning is thus not sufficient to determine the level of consciousness in a given person (A.M Owen et.al. 2006). This is evident from the fact that it is possible for people who emerge from a vegetative state to regain full consciousness, without regaining all of the physical states that are seen a necessary by people who confine to reductive physicalism. Of course this argument is not strong enough to show that there is no correlation between neural and mental activity, in fact it rather seems to lend support from the fact that such correlation on average seems to hold. More critical is the fact that scanning and imaging of the brain does not give any direct insight into the experiential or phenomenal consciousness. They are images of what is going on in the brain, but they really do nothing in terms of explaining how we come to experience through these neural states, and it is not even sure how accurately they can show the neural activity that corresponds to given mental states. For one the activity in a neural system can correspond to mental content, but it is plausible that there is feedback between the different things we are conscious about at the given moment of the screening. Nöe (2009) exemplifies with a discussion of what happens when we are aware of a rhyming sound. At the same time we will both be conscious of the rhyming itself and the words (language) that makes up the rhyme, and as these will intertwine, and "light up" somewhat different areas of the brain, it is not possible to determine exactly which area corresponds to the given

mental state. This will at least be the fact as long as it is impossible to determine whether a person can be conscious of exactly one thing, and exactly what thing this is.

In Nöe's (2009) words, these notions point to the fact that it is not the brain that thinks, we do (a person or another living being), and to explain how consciousness is constituted we need to take more than just the brain into account. One of the reasons why neuroscience has problems with giving an explanation of how consciousness comes to be, and thus bridge the explanatory gap (Chalmers 1995) is the sole focus on the brain, and the assumption that this is where the explanation is to be found. If this is right, it might be a tactical maneuver for cognitive science to take a look at the phenomenological/enacted way of thinking about and explaining consciousness. Alva Nöe's (2009) main point here is that to get a clearer glimpse of consciousness, and how it actually can be explained, something more than just the brain needs to be taken into account in order to find out what the actual constitution of consciousness is.

But if the brain is not what (at least not on its own) constitutes consciousness, then what is? And how is this alternative explanation for the constitution of the mind able to give a better understanding and explanation of consciousness? What is the actual role of the brain seen from this perspective? In other words, what constitutes consciousness according to Nöe and the enacted theory of mind?

As we have seen, the question goes from being "How does consciousness arise in the brain?" to something like "How does life, and the interaction between brain, body and the phenomenal world together give rise to consciousness?". Nöe, along with Susan Hurley (2003), have pointed to the fact that a somewhat artificial boundary is assumed between the brain and the rest of the world, including the body. The fact is rather that consciousness is not limited to the nerve cells; the nerve system or even the brain, there is no actual boundary here to stop us from incorporating the rest of the body, or even the outside world in the discussion of how consciousness comes to be. Nöe (2009) refers to situations in which the neural systems of the brain do not function as they do in ordinary living beings to show how it is actually necessary to incorporate the way we are situated in the world as living beings with a body. By appealing to the plasticity of the brain, and the way a young brain quickly adjusts and adopts to its environment, Nöe shows how the neural activity itself depends on interaction, rather than just input from the surrounding world in order to function as it does. This adaption happens through structural wiring in the brain, and as the brain keeps at least some of its plasticity over the years the structures can probably be altered throughout a

person's life. So, as the way we relate to the world change, the structures of physical activity in the brain changes with it. An example related to this is found in an empirical study of the brain functions of ferrets (Sur, Angelucci and Sharma 1999). In an experiment the scientists blindfolded newborn ferrets to see whether this would affect the later development of vision. The result was that when eyes of the ferrets later were uncovered they could still not see, and the ability to develop sight was no longer present. This shows that a young brain is susceptible to influence from the outside world, and that the kind of interaction that is undertaken early in life is crucial for how the brain develops. The brain, and with it the kind of mental states and conscious experience we can have is decided by how we interact with the world. Speaking of human beings, the way we are socialized, by culture, language and relationships to others can all play a role for both the structure of the brain, and in turn the way we come to experience the world. Nöe also points to the fact that individual brain cells or even areas in the brain does not have an intrinsic quality that makes them responsible for giving us the conscious content that is our experience. A study from MIT (Nöe, O'Regan 2001) shows that if the neural systems of a brain is rewired so that the neural pathway from the eye leads to the auditory cortex, and the pathway from the auditory organ winds up in the visual cortex, this does not make the animal hear with its eyes and see with its ears. Instead the auditory cortex takes over the job of the visual cortex and vice versa. According to Nöe this kind of plasticity shows how the neurons themselves are not responsible for how the world shows up in consciousness, and that the way our body (here eyes and ears) interacts with the world has something to say for the brain works as well.

Even further, studies of tactile-visual substitution have showed that using an apparatus that substitutes visual input with tactile stimuli that correspond to the image of the participants visual field (Bach-y-Rita 1972). The apparatus, engineered by Paul Bach-y-Rita in the 1960's (Nöe 2009), uses a camera that is connected to an instrument able to submit tactile stimuli somehow corresponding to the images recorded by the camera. Remarkably this apparatus made it possible for visually impaired people to "see" what was going on in front of them, and they were able to pick up or point to things, as well as hitting a ball with a racket. Like in the case of the rewiring mentioned above, we see here that it is possible for a specific type of sensory input to be translated in the brain so that it gives a different kind of experience. Even though the sensory input that is given by the sensory

substitution apparatus is tactile, a person who is adapted[12] to the signaling of it will experience not only touch, but also a kind of seeing. The somatosensory cortex of the brain that receives the signals does in other words alter the way the signals are experienced. Theories that base much of their evidence on finding neural correlates of consciousness is thus in for a difficult task. If there is no intrinsic and given way that a neural system realizes conscious experience, it would seem like a bad idea to look for the constitution of consciousness only here, and instead of adopting this identity theory hypothesis one could look at how the world relates to us as involved beings. Neurons in the brain react to the way we relate to the world, and make it possible for the system to bring about the experience corresponding to our activity, the world is thus of an utmost importance to the way consciousness is being facilitated in the brain. Interaction with the world does in other words have a constitutive effect on the brain, here realizing tactile input as visual experience.

Empirical studies like the ones employed by Alva Nöe shows how the mind is related to the body and the world. But this in itself is not revolutionary thinking. In cognitive science and traditional analytic philosophy the fact that body, world and mind interact is apparent in the because we have a perception of the world we live in, made available through the senses along with the neural system. The fact that we have such sensory input is seen as crucial for the content of the mental states, as it directs the way we behave in, and relate to the surrounding world. It is after all the fact that we get some sort of input from the "outside" through our sensory organs, and this input does give some sort of direction to the way we behave. In other words there are causal effects from the sensory input to the behavioral output. The influence from the outside world alter the firing of neurons, and thus the structures of the brain, pointing to a definitive causal effect from the world as it is given through our senses to the physical states of the brain. The fact that empirical studies show that the way our bodies are related to the world have an impact of the way the brain looks and behave, does in other words have a causal relevance for the way physical reductionism (mind/brain identity theory) relates the world to consciousness. And from what I have mentioned earlier this point is also important for Nöe in his enacted theory. But, for Nöe, this relationship between the world and consciousness is not a purely causal relation, as the way consciousness is directed at the world is also seen as constitutionally relevant for

[12] This kind of adaption is very fast and can be done in minutes and hours rather than days and weeks (Nöe 2009 p. 57)

consciousness. So the question is, how does Nöe get from this purely causal relationship to thinking that the body and the world actually are constitutional for consciousness?

The notion mentioned above points to a problem in Nöe's enacted approach. The moment of critique is that the enacted explanation of consciousness is inclined to some confusion between causal and constitutive relations. Some of the confusion stem from the way Nöe views the relationship between the world and the mind, as he is reluctant to think that the world can somehow be torn away from the way we talk about consciousness. But if the world in itself is only causally relevant for the way we are conscious, then there is a chance that Nose way of viewing the constitution of consciousness is based on a misunderstanding, or confusion of what is to be understood as causally and constitutionally relevant for consciousness. The problem is apparent in the way Nöe uses empirical examples to illustrate how the brain and mental states are connected to the world and the body. The examples and connections he makes use of can most of the time easily be referred to as causally relevant for an explanation of consciousness, and this is the way his examples are used by cognitive scientists and analytical philosophers of mind. Nöe on the other hand uses the examples to point to how the world (and body) are crucial for the fact that we can have any conscious experience at all, and as such he use the empirical examples to point out how the relation between the world and "the self" actually will have to be described as a constitutional factor for consciousness. A reason for this might be that Nöe and the enacted theory finds it difficult, or even impossible to separate the brain from the world and the body. This notion is to some extent illustrated by the brain in a vat thought experiment that I mentioned earlier, and the fact remains that Nöe (2009), along with other phenomenological thinkers interpret the evidence given by the type of empirical studies I have discussed here differently than a cognitive scientist. In the next section I will argue that there is at least one way in which it is possible to understand the world as a constitutional factor for consciousness. This will also lead to an investigation of different readings of Nöe, and how I think these different readings can be used in relation to the Hard Problem. Even so, what we have in front of us here is an objection to how the enacted theory accounts for the constitution of consciousness. That is, if the arguments meant to support the thought that consciousness is constituted by the relation between brain, body and world is resting on a confusion between what is causally and constitutionally relevant, one of the motivations for endorsing it is undermined.

4.2 A New Understanding of the Constitution of Consciousness

Still, the way phenomenology accounts for consciousness might open up for some new ways to understand the role of the body and world in relation to the brain and the mind. For it is still possible to think about the explanation for why we are conscious as a result of the way we are related to the world, even though it is difficult to find evidence for it in the strictly analytical way. One way of viewing this is that it is impossible to separate the conscious mind and experience from the way neural activity realizes consciousness. If this is true, the notion naturally leads to a dynamical understanding of consciousness, where a constitutive explanation refers to a dynamical process that links neuronal structures and the world together. This dynamical process refers to the way body and mind is related to and intended towards the world. The relation can be seen as a dynamical process because it is no longer a one way cause to effect relationship, where the environment is seen as the causal factor that affects the physical states in the brain, and as such also the experiential, intrinsic mental states. Instead the dynamical understanding of consciousness lets the mind itself take an active part in the way the world is constituted in our experience of it. The point is that the world does not "show up" as long as we do not relate to it, the way the world is perceived by us depends on our actions in it. The structure of consciousness should be understood as something more than a receiver of sensory input and a computational processor of information. It should rather be seen as a partaker, as experience first show up when we relate to, and act in the world. When consciousness, and the relationship between the world and mental states is described this way, it opens up for an interpretation where intrinsic (subjective) aspect of conscious experience is disclosed in the extrinsic, or observable aspects of how we relate to the world.

So, even though the world might be seen as causally linked to the mind in an empirically detectable way, I will try to show in what follows that this does not necessarily exclude the world as a constitutionally relevant, depending on how we read and interpret Nöe's enacted theory.

If we are to believe Nöe, the plasticity of the brain shows how the outer world makes a contribution to the development of consciousness, as the neural structure does not itself contain all the necessary facts about how we become conscious. As we see in the example

of the sensory substitution experiment, there can be an alteration of experiential content without the occurrence of any large scale change in the structure of neuronal patterns in the brain (as there is no time to make significant changes or rewiring in the short time it takes for a person to adapt to the tactile form of vision). And if the structure of neural activity is not all that is necessary for the constitution of consciousness, it makes little sense to only map out the brain on the search for neural correlates of consciousness. Instead we would also necessarily have to take a look at how the mind relates to, or are intended towards the world if this too is a constitutionally important factor. So according to Nöe consciousness is an enactment that happens in our meeting with the world. When the body is in the world we have to adapt to it, and use our bodies when we relate to it, and this interaction, or rather enactment is what constitutes consciousness, with the brain and its neural activity serving as a facilitator for the dynamic pattern between mind, body and brain (Nöe 2009, p. 47). So Nöe's enacted approach also incorporates a notion of embodiment, as the way we come to be conscious depends on the fact that we have body as much as it depends on having a brain (Nöe 2012). The role of the brain is in a sense not diminished as it is still what makes the combination of the constitutional parts possible, but it is not the only thing that serves as a constitution for mental phenomena and experience. We need a brain, but this brain is of no use as long as we do not have a body and a world in which this body can be put to use through thoughtful interaction. As such, experience is not something that happens to us, it is something we do, an action seen in the way we relate to the outer world and to other bodies which we perceive to be like our own. That is as an acting and experiencing form of life[13].

So even though Nöe's use of empirical studies might be a misunderstanding as far as it comes to showing how the world is constitutionally relevant for consciousness, it is still possible to interpret the enacted theory as an expression for such a dynamical understanding of the relationship between body, world and mind the way I do here.

The major difference between cognitive/neural science, analytic philosophy of mind and the enacted approach is thus not that the phenomenological theory takes the brain out of the explanation of consciousness, but the fact that analytic philosophy and science

[13] This is also the solution to the problem of other minds, a classical problem in analytic philosophy of mind that asks the question of how we can have any knowledge of the existence of other conscious beings. The problem arises due to the subjective character of the experiential content of the mind, and works from the hypothesis that it is impossible to get a glimpse of what is happening inside the mind of another being, and thus we cannot be sure that there is anything going on at all. For Nöe on the other hand, other people/beings show up for us as conscious in the way we relate to them in morally based interaction (see Nöe 2009, p. 25-46).

takes the world and, to some extent, the body out of the equation. In the enacted theory of mind the brain does play a crucial role as a facilitator for the occurrence of consciousness, but what is seen as really constitutional is the body that relates to the world as an intentional being (Heidegger 2007). In classical phenomenological thinking, and also in the enacted theory of mind, consciousness is something implicit in the way we relate to, and use the world as place where our lives can unfold. Consciousness becomes a kind of expression for the fact that we are in the world, implying that the world in a sense necessarily shows up for us when we relate to it in a meaningful way. This is a stark contrast to classical analytical thinking of the mind and consciousness, which perceives of consciousness as something indisputably internal, something that happens in a closed "box" out of reach for others. This notion is nicely formulated in Wittgenstein's (1997) famous beetle in a box thought experiment. Here he imagines that all the people in the world have their own designated box containing what everybody refers to as a beetle. But no one can ever open the lid of another person's box, and as such it is impossible to tell whether every box contains the same thing, a "beetle", or in fact if there is really anything in the other's boxes at all (Wittgenstein 1997). If the same is a fact of the mental "content" of the brain, consciousness is indeed something mystical and inscrutable. This is not the case according to phenomenological thinking, where conscious experience is perceived as something that is readily available through the fact that we interrelate and connect to the world with intention (Heidegger 2007). This is a notion that is close to how most people actually perceive and relate to consciousness in everyday life. In the enacted theory the involvement in the world is described as an acting relation, thus making consciousness an enactment, something we do, and not something that happens to us because there is neural activity in the brain. Again we see how the enacted approach sees involvement and our relation to the world as constitutionally relevant for consciousness.

In the light of the discussion of the Hard Problem (Chalmers 1995) of philosophy we see a difference between how the two directions, the phenomenological and the classical analytical, perceive the problem, and how they think it best can be solved. For classical philosophy of mind the problem shows up due to the fact that there is still no completely satisfying explanation for how consciousness is constituted by the brain. I have previously gone over some of the alternatives that have been proposed in order to give an adequate explanation, but there are still serious problems with all of them. One crucial mistake made in most of these theories is that they do not have any serious suggestions in terms of how

the purely experiential factor of mental states are connected with the neural activity of the brain. Many do not even discuss this seriously, and for others the opacity of conscious experience has led thinkers to deny the existence of such states on a whole (Dennett 1991, Churchland 1981). In light of the enacted approach (Nöe 2009) it is, on the other hand, possible to hold that these mental states have an extrinsic aspect, and that consciousness is readily available to those who wish to make inquiries on the subject. The intrinsic and experiential qualities of the mental states are in other words visible (extrinsic) in the way we relate to the world.

This reading of the enacted approach leads to an interpretation where it is possible to look at the Hard Problem of consciousness as something that does not show up in the way it does for classical neuroscience and philosophy of mind. It is clear in phenomenology that a constitutive explanation of consciousness should provide an understanding of how the individuals interaction with the world by virtue of being a sensing body gives, or rather shows how the person/being is conscious. In other words, the Hard Problem does not show up for the phenomenological understanding of consciousness at all. But it still seems to me like there is a question connected to exactly how the brain facilitates the occurrence of qualia. If the brain is seen as what somehow enables the conscious experience to appear, it is still necessary to get a better understanding of exactly how this happens. It still seems as though the brain has some crucial role to play, and that this role is not properly explained by the phenomenological enacted theory. The phenomenological thinker might reply that this is a different question, and that it is not crucial to an explanation of consciousness. Instead it should be viewed as a technicality left to the hands of the natural sciences, and not a question for philosophy of mind. Still, phenomenology and the way it opens up for an extrinsic understanding of how experiential consciousness is constituted might open up for making better explanations of this specific aspect of mental states when viewed with the eyes of for example a functionalist, and I will pursue this further in what follows in order to find out whether the enacted theory can be a supplement to functionalism. In addition, the phenomenological account of consciousness as it is stated by Alva Nöe (2009) is also open for the interpretation that there is no real "Hard Problem" for philosophy to solve, this too will be discussed further later in the paper. The third option that I will discuss is whether the intrinsic aspect of consciousness is intertwined with the neural structures of the brain in a way that makes it impossible to separate between intrinsic and extrinsic expressions of conscious experience. If this is the case, and

consciousness is a kind of dynamical process between world and mind, then there is no need for the physicalistic theories of consciousness to hold that the identification of a mental state has to be necessarily extrinsic.

Functionalism is the theory, belonging to the realm of analytic philosophy of mind, that has gotten furthest in the search for an explanation to the problematic notion of qualia (Livingstone 2005). By reducing or identifying mental states with functional states rather than purely physical states in the brain, the theory has gotten far in terms of interpreting how the psychological states and behavior of a being give an expression of experienced consciousness.

By reducing mental states to functional states, functionalism is also able to make sense of (at least to some extent) the problem of mental causation. This problem arise due to the dualistic aspect entailed in the formulation of consciousness as it is used in relation to the Hard Problem. The "what it is like" aspect of consciousness as it is fronted by both Nagel (1974) and Chalmers (1995) involves a view of consciousness that can lay the ground for an epiphenomenal understanding of consciousness (Yablo 1992). If consciousness is understood as an epiphenomenon this leads to a problem related to how we can account for the causal effect from mental states to physical states. This means that if qualia are not reducible to physical states, feelings and intentions such as wanting, wishing and believing can have no causal effect on our (physical) behavior. This view is obviously problematic, as it seems highly unlikely that what we consciously wish to do is irrelevant for our behavior. If mental states can be reduced to physical states, as the physicalistic mind/body identity theory indicates, this problem does not show up as there is no problem of causation between physical entities. But in the identity theory there seem to be no room for qualitative experience either, so the fact that it can provide for an answer to the problem of mental causation is of little help. (Role-) functionalism solves this problem by identifying the mental states with functional states. The trick is that functionalism, in order to avoid dualism and its adherent problems, such as the here mentioned mental causation, is able to reduce consciousness to a physically available level. This kind of reduction is not the same as the physical reduction found in the mind/brain identity theory and role filler-functionalism where mental states are reduced to brain states. But role-functionalism can still be held as a type of reductive theory because of the way mental states are described as functional states. By functionalism consciousness is not understood as an epiphenomenon, but as functional states, and as such the problem of mental causation is avoided. In other words functionalism is not

only successful in terms of avoiding some of the problems associated with physicalism, it is also able to account for the problem connected to dualism and epiphenomenalism.

These functional states are in many cases successful when it comes to saying something about our mental states, as these are often expressed in extrinsically available expressions of behavior (Kim 2005). But according to amongst others Jaegwon Kim (2005), functionalism is still in trouble as far as it comes to accounting for the qualitative aspect of consciousness as the quality of experience is not expressed in the functional states. Kim (2005) exemplifies with the inverted spectrum hypothesis[14] where the way we perceive colors has been altered so that every color looks like its opposite (red looks green, and so on). According to Kim, the qualitative aspect of the colors are really irrelevant for the functional state, as long as the person with the inverted color spectrum have the same relationship to, and ability to discriminate between the colors as a person with "normal" color vision. For example, this means that if a person with an inverted color spectrum perceives grass and green salad as having the same color, a color that is opposite of ripe tomatoes (and so on), there is no difference for functionalism, and the theory can still make sense of the behavioral response and use it to determine a functional state. The qualitative experience is, on the other hand, not available in the functional state, and the way a color (or a wish for that matter) feels to us in our subjective way of experiencing it is still without causal power, and is thus unable to account for how we behave. According to Kim (2005) the intrinsic nature of qualia is what makes it unable to be identified with a functional state. In result, further work is needed in order to make functionalism able to account for constitution and rise of consciousness.

If we take Kim's (2005) suspicion that qualia, the distinct feelings of an experience, are impossible to "functionalize" due to their intrinsic character (i.e., they cannot serve as functional states because they are not available from a third person point of view, and cannot be put into a scientific system) the adoption of the enacted theory of consciousness might serve as a way to make consciousness open for scientific description. The point is that the enacted theory of consciousness characterizes consciousness as an extrinsic factor of human life, rather than something that is invisible and closed within a wittgensteinian beetle box. Consciousness is in other words understood as something that lies open in the day, and not something that is purely subjective and hidden from the third person perspective. By describing consciousness as something that happens "outside" the brain, in

[14] See Chapter 2

the meeting between a subject and the world, it is possible to say something objective about conscious experience, as this experiential content is apparent in the way we relate to and use the world in a meaningful way. If Nöe and other phenomenological thinkers are right about this, the extrinsic quality of conscious experience can be used to individuate and describe functional states, as they are open for empirical description and research.

This thought, that the intrinsic, experiential qualities of consciousness can be seen as expressed in the way we relate to the world has led to a notion about the naturalizability of phenomenology. The project is undertaken by amongst others Petitot, Varela, Thompson and Nöe himself in the book "Naturalizing Phenomenology" (1999), and is directed towards making the phenomenological understanding of the world available to empirical study through various experiments and examples. Success on this field might plausibly make it easier to point out the "intrinsic" states that are necessary to supplement functionalism, even though the project collides somewhat with Husserl's (1913) own view, that phenomenology should not be naturalized due to its a priori character, or the fact that both the world and we, as conscious beings are something given, just by being in the world. If, on the other hand, the project is successful, as Nöe's enacted theory hopes to be, this could provide a logical frame-work for functionalism in which conscious experience (qualia) can be referred to as functional states, enabling it to give a functionalistic explanation of the relationship between the body/brain and consciousness.

To summarize, my thought is that there are several ways to read the phenomenological enacted theory of mind. The first one I have discussed investigates an interpretation of the enacted approach where the intrinsic, experiential states are not only intrinsic and experienced by the subject, but also available to a third person point of view, made visible in the way we relate to the world. This reading depends on Nöe's argument (2009, 2012) about how the world and the body, or rather how the body acts within this world plays a crucial role for how the mind is constituted. For Nöe it is nothing less than a fact that through this relation we enact our own consciousness, and this enactment or action is in my reading of Nöe taken to be constitutional for consciousness. If this is true, then consciousness, as long as the term is used to include experiential and qualitative aspects, can be understood as something that is open to all, and not something that is hidden inside of our brains, like Wittgenstein's beetle is locked into its box. This can in some respects be com- pared to how we all, according to Heidegger (2007) relate to the world in a state of moodedness, or with care (sorge) towards our own actions. This moodedness directs the way we relate to the

world, but it is also a reflection of our internal states or experience of the world. When the enacted theory of mind is understood this way it can be linked closely to functionalism, as the problematic intrinsic states now are open for inspection. In other words, the functional states that are expressed in behavior and the relation they bear to sense-input and other mental states can be taken to include the "intrinsic" states that are, according to Haugeland (1978) and Kim (2005), unavailable and out of reach for classical functionalism

4.3 Some Interpretations of Nöe's Approach to Consciousness

The problem with this kind of reading is that it is still unclear exactly how the brain and neural processes are related to the occurrence of consciousness. If one is ready to agree with the statement that consciousness is enactment, and that it is constituted by the way we relate to the world, there is still a problem, or at least unclarity connected to the role of the brain. It seems to me that in the case spelled out here, the brain must at least play some part on the way to consciousness, but no answer, or even speculation is given as to how exactly this happens. In a way the Hard Problem is altered, and is no longer understood as a constitutive question of how consciousness is realized. Rather it is an inquiry of what makes the brain able to give rise to consciousness when it interacts with the body and its actions in the world. This is at least one way to interpret Nöe's enacted approach, and we do, in other words, still not have a full understanding of experiential consciousness. But this kind of reasoning also points to the important problem for traditional philosophy of mind, and especially neuro science, and helps sort some of this problem out.

The problem I refer to is, again, that neuroscience and analytical philosophy has a tendency to look away from, or at least take little notice of consciousness as it appears through the kind of trained introspection performed by phenomenological thinkers (Laughlin 1999). Consciousness itself is clearly a phenomenological concept as it is, for everybody, reached though introspection, as our knowledge of consciousness just because we are conscious beings capable of this kind of introspection. Knowledge about consciousness is thus gained by introspection, and it is thus necessary to use some introspective techniques in order to fully understand it. And a full understanding of consciousness is necessary, as it is

the explanandum for both philosophy of mind and neuroscience. But, as mentioned, this is often overlooked by many who try to explain the phenomenon of consciousness, and explanations tends to lack sufficient substance when it comes to tackle the question of such phenomenal consciousness. In order to make good and correct theories about experiential consciousness, a thorough understanding of the concept is needed, and it is thus clear that any such theory would benefit a great deal from phenomenological and introspective analysis. For example I think much of the reason why philosophers such as Daniel Dennett and others who comply with eliminative theories do not sufficiently take into account this notion of how our understanding of consciousness is made possible. According to this it should be seen as beneficial for neuroscience and analytical philosophy of mind to take phenomenological analysis of consciousness into account in order to make out a good explanation for how consciousness is realized. This is not unlike what is done by phenomenological thinkers such as Nöe, who in order to make a naturalized phenomenology for the sake of explaining consciousness make use of empirical, neuro-scientific experiments and research. The benefit given by this kind of mixture of neurology and phenomenology is that it gives a wider and hopefully more exact way to understand and explain consciousness. Additionally, the problem I mentioned above can be solved by diminishing how we see the constitutional role of the brain. This notion will be further discussed later in this thesis.

Another way to interpret and another possible consequence of Nöe's enacted approach is that the Hard Problem can be denied all together. This thought is quite common for phenomenological thinkers such as for example Shaun Gallagher and Dan Zahavi (2008), and the notion depends largely on an interpretation of how conscious- ness is constituted. In Nöe's enacted theory, and in the phenomenological perspective on consciousness more generally, the constitution of consciousness is found "beyond the brain", in the relation between mind, body and brain, and the way these relate to the world as a place where life can unfold. In the previous section I state that this kind of constitutive understanding can make it possible to view the intrinsic states of experience as really being extrinsic. This corresponds to a dynamical view of how consciousness is related to the world, and the way this view makes it difficult and unnecessary to distinguish between these two kinds of states. Instead they are understood as the same thing, as intrinsic states can be understood to have extrinsic qualities and the other way around. This is because experiential content is visible in the way we relate to the world, and not least to other people. The notion of a state that is only intrinsic can be avoided in this way, as they are now taken to have an extrinsic aspect, presented in the

"outside" world, and expressed in the way we relate to the world. If this is accepted it implies that the Hard Problem of consciousness does not show up for phenomenology, at least not in the same way as it does for an analytic philosopher of mind.

According to Chalmers (1995) the Hard Problem can be defined as the problem of bridging the gap between physical states in the brain and experiential consciousness, and an explanation of this will need to describe the way consciousness is realized by describing the relationship between consciousness and the physical world. This mystery rests on the "assumption" (Nöe 2009) that experiential mental states really are produced or realized by the brain. But if it is room for disagreeing with this premise, it is also possible to argue that there actually are no such difference between the mental and the physical, and as such there is no reason to think that there really is a "hard problem" to be solved. And disagreeing with the premise is exactly what Nöe does. For him it is more relevant how to look at what (causally) changes the structure of the neural circuits of the brain, than how these circuits themselves are the origin of conscious experience, and as such the constitution of consciousness is to be found here, out in the world. So consciousness is constituted by the world, or rather, our relation to it, and the way we act. The notion the consciousness is visible and out in the open is also a clearly intuitive thought, as it reflects the way we relate to other thinking subjects. But when consciousness is perceived as constituted by our (conscious) presence in the world, the brain itself seems to be left out of the equation, as far as it comes to being an essential factor for consciousness. It is seen as an organ, along with all the others, necessary for us to be alive. The role of the brain is to process sense data, and makes us able to relate to the world, but its role is limited to purely causal matters. The question of how consciousness is realized by brain in this relation does not arise simply because consciousness is this relation itself. Consciousness is in other words the mind and body (with brain) that relates to and acts in the world.

According to this interpretation then, the Hard Problem just does not show up. As such the enacted approach certainly "helps" to solve the hard problem of consciousness; by denying that the problem is even there.[15] This notion seems to be questionable for many

[15] This notion bears an interesting similarity to eliminative theories of consciousness, even though this kind of theory must on a whole be regarded as the diametrically opposite of the phenomenological account. The similarity stems from the fact that the eliminative approach denies the existence of the hard problem of consciousness, just as many phenomenological positions do. And surprisingly it seems like they both arrive at this account due to a notion that there are actually no fundamental difference between mental states and physical states. But here stops the similarities. For while phenomenology sees consciousness as the way we are physically present in the world, something that thus has no mystical origin in physical states in the brain. Here the world is consciousness, and consciousness is the world. In other words there is no difference

reasons. One of them is the method by which the conclusion is derived, as this seems to be somewhat distanced from scientific and empirical studies, instead it is more of a folk psychological type of theory and a demystification of consciousness. This might be a good thing when it is used to provide a structure for what the exact object/phenomenon of study is, but for a scientist, or an analytic philosopher for that matter, it would seem close to no explanation at all, as it is not based on facts that are scientifically verifiable. Rather, phenomenology lends its support purely from the kind of "skilled" introspection that is typical for phenomenological philosophy. Alva Nöe tries to use a naturalized kind of phenomenology on order to make an empirical foundation for his theory, but these are all identifications of causal links between the way the living subject relates to the world, and as such a constitutive explanation is not really available through the kind of evidence he makes use of. On the other hand this reading or interpretation might be too strong, and Nöe's enacted approach should not necessarily be seen as an effort to deny the existence of a relation between brain and conscious experience, or the existence of an intrinsic aspect in experience. Even so, it could still be seen as an independent effort to find an answer to the Hard Problem of consciousness, and thus an explanation of the mind/body problem. Even so, if we disagree with the notion that the explanation can stand alone as an explanation of consciousness, Nöe's approach to the philosophy of mind can, beneficially, be seen as an effort to change the way we regard the boundaries between the physical and the mental. A better interpretation of this is that Nöe suspends some of the boundaries between the physical and the mental, and the intrinsic and extrinsic aspects of conscious states in a way that gives a better understanding of consciousness in favor of a more coherent and sufficient explanation of the relation between mind and brain.

This last reading of Nöe is better in terms of portraying the suggestions actually found in the enacted approach, as it relates to the dynamical description of the relationship between consciousness, brain and our bodily relation to the world. In addition it seems to be the best in terms of being useful to better the understanding of the dynamical character of the relationship between these "parts". This dynamical character implies that the lines drawn out by analytic philosophy of mind and cognitive (neuro) science are somewhat artificial.

---

between the physical word out there and the way we experience it consciously (Husserl 2012). A line drawn between the two would thus be artificial. Eliminative theories on the other hand denies the existence of an explanatory gap on account that there are no such thing as conscious experience (Dennett 1991). Mental states are nothing over and above the physical states of the brain. In a way we are machines, and neural structures are the components of the machinery, and consciousness understood as subjective and intrinsic experience cannot be explained by this system, and so have no place in it.

The lines that are drawn between body (brain) and mind, between the subjective and the objective, and between the world and the person, in many ways echoing the words of Descartes meditations (2013) and the dualism we find here. This kind of dualism is found in the boundaries mentioned above, and is by all means one of the main reasons for the occurrence of the Hard Problem. The dynamical description, or understanding of the relationship between perception, bodily involvement and awareness and conscious experience opens up for a different understanding, and, to a large degree eliminates these boundaries. By appealing to the process by which we become conscious as an interaction between all of them. Mental processes are as such not understood as processes in the brain alone, but are described as a process that stretches out, and incorporates both the body and the surrounding environment. In other word we should not draw a line between intrinsic states and corresponding extrinsic states, or between the fact that we are living beings (with a brain) and consciousness, they are as two sides of the same thing, and are intertwined with each other in a way that makes it impossible to separate them in terms of attributing the fact that we have conscious experience to one rather than the other.

This notion, that the line traditional analytic philosophy of mind, along with cognitive neuroscience draws between the intrinsic and extrinsic gives a somewhat artificial understanding of consciousness can be used to create a dynamical form of explanation, in an effort to find a solution to the hard problem. This kind of approach takes into account both the phenomenological description of how conscious- ness "work", in terms of how the mind is structured in order for experiential content to be present, along with the progressive work found in modern neuroscience on the relation between consciousness and the brain. The premise for this kind of research is that there is actually no real, static boundary between consciousness and the world, or between extrinsic and intrinsic "parts" of mental content. These two ways of looking at consciousness does not really belong in different spheres, but are intertwined and work together in order to realize the experiential content of the mind. On this account it is natural to make use of both the phenomenological approach to consciousness, as well as the knowledge we have of the relation between brain and mind from neuroscience.

This understanding of how consciousness is embodied, and dependent on a relation to the world can yield different kinds of theories or efforts to merge the phenomenological understanding of consciousness as embodied and related to the world, and explanatory models that are closer to those of natural science. Tononi (2000), as I mentioned above, uses a

description of the dynamical structure of consciousness as it is described here[16] as a way to gain a better understanding of the structures that underlie consciousness. The point is to create a phenomenological account or analysis of consciousness in order to lay the ground for a neuro-scientific explanation of consciousness. Again the point is that consciousness needs to be properly understood in itself, for the sake of giving an adequate explanation[17] of its occurrence or realization.

So the dynamical understanding of consciousness can be used as a more steady foundation for explanations of consciousness, and works as a neat way for clearing out the kind of explanation that would do the job of accounting for the realization of consciousness. Or at least it works well to weed out those explanations that do not give adequate results in terms of explaining the structural features of consciousness as they are described by phenomenology. This can, as mentioned before be seen in relation to a role filler kind of functionalism, as the phenomenological description of consciousness is seen, and used as a functional state that points to the (underlying) neuronal states, as they are, or at least should be explained by modern neuroscience. This again can be used to point to how there is quite a strong connection between phenomenology, here understood as embodied, dynamical theories, and functionalism. When used as by Tononi (2000) and other similar thinkers, it is also possible to draw a line to role filler functionalism, while the kind of work done by Nöe (2009) and amongst other Susan Hurley are better understood in relation to role functionalism, where the functional states themselves are found in the dynamical interaction between the environment and the embodied brain. These are not the words of Nöe himself, but it is striking how well the two fits together. In role functionalism it is seen as unnecessary to point to neural states in order to determine and denote mental states, and instead functional states do the job. As such the constitution of consciousness is found in the way sensory input, different (other) mental states and the behavioral output are related, and if functionalism can be supplemented with the enacted approach for the sake of also being able to account for the seemingly intrinsic, felt qualities of experience, it is a promising solution to the hard problem of consciousness. As far as I can see this should be seen as a logical extension of role functionalism as long as Nöe's approach is held as giving a

[16] Though Tononi's strategy is based on a wider understanding of phenomenology, and does not use Nöe's enacted approach directly, the dynamical understanding of consciousness is present in both theories.
[17] For more information about Tononi's project, see for example "A universe of consciousness : how matter becomes imagination" (2000).

correct description of the structures and constitution of consciousness, and hopefully this can be used in order to solve the Hard Problem of the relation between body and mind.

One of the reasons for the good fit between functionalism and enacted/embodied phenomenology is that they are both in a way reductive theories, as they both see consciousness as something that can be explained by an underlying structure or phenomenon. In other words consciousness can be reduced to, and explained in terms of the dynamical process between perception, embodiment, and the way we are related to the world. Consciousness, with its experiential features is understood as some- thing factual, and is possible through the way we are situated in the world as this gives us something to be conscious of. But in the case of Nöe (and role functionalism) consciousness is not (necessarily) reducible to physical structures in the brain.

One of the important points in Nöe's enacted approach is that there is no point in looking for neural correlates of consciousness in the brain, as a strategy for explaining consciousness. The constitution of consciousness is not to be found here, but should rather be understood as an expression for the way we are situated in the world as acting subjects. The role of the brain is important because it facilitates a way in which body, brain and world are able to interact, and consciousness itself is constituted by this relation. This understanding is compatible with functionalism, and can to even be seen as a functionalistic kind of explanation on its own, if we refer to the enactment of consciousness as a functional state. So maybe Nöe is right in thinking that it is unnecessary, and even impossible to find the constitution of consciousness when the scope of the search is limited to the brain, and that there is actually no point at all to point to a specific neuronal correlate to consciousness in the explanation, as the constitution is not to be found here at all. And when this is supplemented with role functionalism, it can make sense also within an analytically angled tradition of philosophy of mind. My conclusion is thus that there are a way in which to read, and interpret Nöe that works well within an analytical perspective on philosophy of mind. Just by changing how the constitution of consciousness is understood, by incorporating the way we relate to the world as subjects instead of passive, receiving objects, much can be done in relation to answering the Hard Problem in philosophy of mind. At least it goes to show that there is more to consciousness than what we seeing when we look inside the head and study the brain.

5 Conclusion

In this thesis I have discussed a way of supplementing functionalism with a modified reading of Alva Nöe's enacted approach. The conclusion I propose is that, since role-functionalism allows for consciousness to be identified with functional states, it should be possible, with a phenomenologically inspired understanding of conscious- ness, to interpret the functional states as also comprising the phenomenological aspects of subjective experience. In this way consciousness, with the experiential aspect of qualia can be accounted for with a supplemented functionalism.

I have discussed the distinction between role filler-functionalism and role-functionalism, as these two interpret the relationship between consciousness and functional states differently. I also propose two different ways in which phenomenology can be used to supplement these functionalist theories.

First, role-functionalism holds that consciousness can be identified with the functional states themselves, but the theory runs into some problems in the effort to answer the hard problem because subjective experience is not part of the way the functional states are described. This is because the functional states are classically used to refer to the relationship between sensory input, behavioral output and other relevant mental states. The problem for role-functionalism in relation to conscious- ness as it is understood in the Hard Problem is that these functional states do not entail experiential consciousness. Rather they are comprised of empirically observable instances, or more accurately the relationship between sensory input, behavior and other relevant metal states. We could in other words imagine one of Chalmers' zombies having the relevant functional state, even though it lacks conscious experience (Chalmers 1996). But by supplementing functionalism with an enacted understanding of consciousness, it might be possible to describe functional states as containing an expression for experience and subjective consciousness. This is because the subjective aspect of experience can be seen as something that is actually visible in the way we relate to the world according to the enacted theory.

In the case of role filler-functionalism the functional states are understood differently, as they are rather used as pointers out the underlying physical states responsible for the realization of consciousness. The job for the functional states is thus to point to, or denote the physical states responsible for the mental states. My conclusion here is that it is

possible to make use of the phenomenological description as a starting point for a neuro-scientific explanation. The point is that phenomenology offers a thorough description of the way the mind is and of how it operates in relation to the world, and this gives a solid ground on which a physiological theory of mind can be built. This is similar to what is done by Tononi (2000) and the supporters of the Canberra plan. These kinds of theories see it as a mission to uncover the underlying, or a priori structures of consciousness in order to create an operational theory of consciousness and experience. The way these two are connected to phenomenology is that they use the outlines found here as a premise for how consciousness should be explained. By setting this up as a starting point, both Tononi and the supporters of the Canberra plan get a clearer understanding of how a physical explanation must be in order for the conscious aspects described by phenomenology to be realized.

I have also pointed to another suggestion, where the answer to the Hard Problem is to explain the whole thing away. This conclusion depends on a specific reading where the constitution of consciousness is understood as visible in the way we are present in the world; as conscious and acting subjects. If consciousness is understood in this way, as something implicit in the way we are in and relate to the world, the Hard Problem of answering the question of consciousness does not even arise. A critique of this view involves an attitude where this explanation is understood as both unclear and unscientific, and I think phenomenology has a better chance of gaining support as a supplement to analytic philosophy rather than as an independent alternative.

These conclusions have been reached through a discussion of several approaches to how the Hard Problem can be answered, and I have especially focused on three ways to understand the Hard Problem. For Chalmers (1995) the problem shows up due to the subjective character of mental states (consciousness), and the fact that these are out of reach of ordinary science, because they are not empirically accessible through an investigation of the brain. It is, on the other hand, possible to find correlations be- tween the different physical states of the brain and the reported conscious experience of a subject, but as these are merely correlations, they do not in themselves have any explanatory value. This fact that there is no evidence for subjective consciousness to be found in the brain is used by some supporters of reductive physicalism such as Daniel Dennett (1991), to maintain that as long as there is no evidence for experiential consciousness to be found in the brain (above the mentioned correlation), we really have no reason to think that consciousness exists at all.

If this is the case, then there is no reason to even talk about a Hard Problem in relation to consciousness, as a thing that does not exist is in no need of an explanation.

This thought is similar to the reading of the enacted theory mentioned above, but apart from this the two theories differ significantly. In my opinion, the fact that phenomenology takes the notion of consciousness seriously, and seeks to account for how it actually is to be in a state of consciousness, is exactly what gives this theory an advantage in relation to the Hard Problem. An advantage that is evident from how phenomenology and the enacted approach are used in this thesis, as a supplement to analytic philosophy and functionalism. For by widening the perception of the constitution of consciousness, it is possible to explain the experiential content of mental states in a way that relates to the physical description of the world. This notion can in turn be beneficial in the effort to explain the Hard Problem of philosophy of mind.

List of References

• Armstrong, D. (1968*), A Materialist Theory of the Mind,* London: Routledge and Kegan Paul.

• Bach-y-Rita, P. (1972), *Brain Mechanisms in Sensory Substitution*, New York: Academic Press.

• Bechtel, W. and Mundale, J. (1997), *Multiple Realizability Revisited*, from the Proceedings of the Australian Cognitive Science Society, available from: http://mechanism.ucsd.edu/research/multiple.htm

• Block, N. (1980*), Readings in the Philosophy of Psychology*, Volumes 1 and 2, Cambridge: Harvard University Press.

• Block, N. (1994), *Qualia*, In Samuel Guttenplan, ed., A Companion to the Philosophy of Mind, 514-520. Oxford: Blackwell Publishers.

• Block, N. (2006), *Wittgenstein and Qualia*, NYU, Available from http://www.nyu.edu/gsas/dept/philo/faculty/block/papers/Block_Putnam.pdf

• Chalmers, D. (1995), *Facing up to the Problem of Consciousness,* in Journal of Consciousness Studies 2: 200-19.

• Chalmers, D. (1996), *The Conscious Mind: In Search of a Fundamental Theory*, Oxford: Oxford University Press.

• Chalmers, D. (2002*), Does Conceivability Entail Possibility?*, in Gendler and Hawthorne 2002, 145–200.

• Chalmers, D. (2006), *Strong and Weak Emergence*, In (P. Clayton and P. Davies, eds) The Re-Emergence of Emergence, Oxford: Oxford University Press.

• Churchland, P. M. (1981), *Eliminative Materialism and the Propositional Attitudes*, Journal of Philosophy 78: 67–90.

• Craver, C. (2007), *Explaining the brain - Mechanisms and the Mosaic Unity of Neuroscience*, Oxford: Clarendon Press.

• Crick, F. & Koch, C. A. (2003), *A framework for Consciousness*, Nature Publishing Group, Available from http://codatest4.library.caltech.edu/26/1/438.pdf

• Dennett, D. (1978), *Brainstorms: Philosophical Essays on Mind and Psychology*, Bradford Books.

• Dennett, D. (1988*), Quining Qualia*, in A. Marcel and E. Bisiach (eds.), Consciousness in Contemporary Science, 1988, new York: Oxford University Press: 43–77.

• Dennett, D. (1991), *Consciousness Explained*, Boston : Little, Brown and Co

• Descartes R. (2012), *Meditations on first philosophy*, edited and translated by John Cottingham, Cambridge, England : Cambridge University Press

• Feigl, H. (1958*), The 'Mental' and the 'Physical'*, in Feigl, H., Scriven, M. and Maxwell, G. , Minnesota Studies in the Philosophy of Science: Concepts, Theories, and the Mind-Body Problem, Minneapolis : Minnesota University Press.

• Fodor, J. (1968), *Psychological Explanation*, New York: Random House.

• Gallagher, S. and Zahavi, D. (2008), *The Phenomenological Mind*, Abingdon, Oxon: Routledge.

• Geroch, R. (1978), *General Relativity from A to B*, Chicago: The university of Chicago Press.

• Haugeland, J. (1978), *The Nature and Plausibility of Cognitivism*, in Behavioral and Brain Sciences 2: 215-260.

• Heidegger, M. (2007), *Væren og Tid*, Oslo: Pax.

• Hume, D. (2010) *Enquiry concerning Human Understanding*, Boston: MobileReference.com

• Hurley, S. and Nöe, A. (2003*), Neural Plasticity and Consciousness*, in Biology and Philosophy 18: 131-168.

• Husserl, E. (2012), *Ideas: general introduction to pure phenomenology*, Abingdon, Oxon: Routledge.

• Kim, J. (1989), *Mechanism, Purpose, and Explanatory Exclusion*, in J. Kim, Supervenience and Mind, Cambridge: Cambridge University Press.

• Kim, J. (2005), *Physicalism, or Something Close Enough*, Princeton and Oxford: Princeton University Press.

• Kripke, S. (1980*), Naming and Necessity*, Cambridge, Massachusetts: Harvard University Press.

• Laughlin, C. D. (1999), *Biogenetic Structural Theory and the Neurophenomenology of Consciousness*, in Toward a Science of Consciousness 3, Cambridge: MIT Press.

• Leibniz, G.W. (1989), *Philosophical Essays*, translated by Roger Ariew and Daniel Garber. Indianapolis: Hackett.

• Levine, J. (2001), *Purple Haze*, New York: Oxford University Press.

• Lewis, D. (1970), *How to Define Theoretical Terms*, Journal of Philosophy, 67: 427-46.

• Livingston, P. (2005), *Functionalism and Logical Analysis*, in Phenomenology and Philosophy of Mind, New York, Oxford: Oxford University Press.

• Lowe, E. J. (1992), *The problem of psychophysical causation*, Australasian Journal of Philosophy 70, 263–76.

• McLaughlin, B. P. (2006), *Is Role-Functionalism Committed to Epiphenomenalism?*, in Journal of Consciousness studies, Volume 13, Numbers 1-2, pp 39-66, Imprint Academic.

• Merleau-Ponty, M. (2012), *Phenomenology of perception*, Abingdon: Routledge.

• Nagel, T. (1974), *What is it like to be a Bat?*, In Philosophical Review 83: 435-456.

• Nöe, A. and O'Regan, K. (2001), *A sensorimotor Account of Vision and Visual Consciousness*, in Behavioral and Brain Sciences 2001, 24: 939–1031.

• Nöe, A. (2009), *Out of Our Heads*, New York: Hill and Wang.

• Nöe, A. (2012), *Varieties of Presence*, Cambridge: Harvard University Press.

• Owen, A. M. et. al. (2006), *Detecting Awareness in the Vegetative State*, in Science 8 September 2006: Vol. 313 no. 5792 p. 1402.

• Petitot, J. et. al. (1999), *Naturalizing Phenomenology*, Stanford: Stanford University Press.

• Place, U.T., (1956), *Is Consciousness a Brain Process?*, in British Journal of Psychology, 47: 44–50.

• Putnam, H. (1967), *The Nature of Mental States*, reprinted in Putnam (1975), Mind, Language, and Reality, Cambridge: Cambridge University Press.

• Sartre, J. P. (1957), *Being and nothingness: an essay on phenomenological ontology*, London : Methuen.

• Smart, J.J.C., (1959), *Sensations and Brain Processes*, Philosophical Review, 68: 141–156.

• Sur, M., Angelucci, A. and Sharma, J. (1999), *Rewiring the Cortex: the role of patterned activity in development and plasticity of neocortical circuits*, in Journal of Neurobiology 41, no. 1: 33-43.

• Sylwester, R. (1995), *A celebration of neurons: an educator's guide to the human brain*, Alexandria, Va. : Association for Supervision and Curriculum Development.

• Thompson, E., Pessoa, L. and Nöe, A. (1999), *Perceptual Completion: A Case Study in Phenomenology and Cognitive Science*, in Naturalizing Phenomenology, Stanford: Stanford University Press.

• Tononi, G. and Edelman, G. M. (2000), *A universe of consciousness : how matter becomes imagination*, New York : Basic Books.

• Van Fraassen, B. C. (1980), *The scientific image*, Oxford : Clarendon.

• Yablo, S. (1992), *Mental Causality*, The philosophical Review, Volume 101, Issue 2, P. 245-280

• Ylikoski, P. (2013), *Causal and Constitutive Explanation Compared*, Published online by Springer Science - http://link.springer.com/article/10.1007%2Fs10670-013-9513-9

• Zimmerman, D. W. (2004), *Should a Christian be a mind-body dualist?*, In Contemporary Debates in Philosophy of Religion, Blackwell, Malden MA.