Universidade do Vale do Rio dos Sinos
School of Humanities
Graduate department of Philosophy
PhD Program in Philosophy

Bianca Andrade

FROM CONSCIOUSNESS TO NOTHING

São Leopoldo, 2021.

BIANCA ANDRADE

FROM CONSCIOUSNESS TO NOTHING

PhD dissertation submitted to the Graduate Program of Philosophy to obtain a PhD/Doctoral Degree in Philosophy – Ethical Systems.

SUPERVISOR: MARCO ANTONIO AZEVEDO, PhD, MD.

CO-SUPERVISOR: ZOE DRAYSON, PhD.

A553f Andrade, Bianca.
From consciousness to nothing. / Bianca Andrade – 2021.
91 f. : il. ; 30 cm.

Tese (doutorado) – Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Filosofia, 2021.
“Orientador: Prof. Dr. Marco Antonio Azevedo
Coorientador: Prof. Dr. Zoe Drayson”

1. Bioética. 2. Consciência. 3. Distúrbios de consciência. 4. Filosofia da mente. I. Título.
CDU 159.9

Dados Internacionais de Catalogação na Publicação (CIP)
(Silvana Dornelles Studzinski – CRB 10/2524)

PhD dissertation submitted to the Graduate Program of Philosophy (Ethical Systems subdivision) of University of Vale do Rio dos Sinos in collaboration with the University of California and with sponsorship from CAPES in order to obtain a PhD/Doctoral Degree in Philosophy – Ethical Systems.

To my loving husband, aunts and parents for the financial and emotional support during my academic journey.

To my son Mateus, born within the making of this dissertation, and to my daughter Clarice, who is due a couple of months after my PhD examination board.

To my supervisors Professor Marco Azevedo and Professor Zoe Drayson for the countless hours of support in order to make this PhD dissertation possible.

“Consciousness, unprovable by scientific standards, is forever, then, the impossible phantom in the predictable biologic machine, and your every thought a genuine supernatural event. Your every thought is a ghost, dancing.”
— Alan Moore, Promethea, Vol. 5

ABSTRACT

What happens when consciousness fails to reach its full and supposedly “higher” level? How much consciousness must one have in order to be recognized as being “conscious”? Could we say that if an individual does not meet a basic standard of conscious activity then we cannot take her as a conscious being? What is sufficient to claim someone to be conscious, and how do we address the so-called “disorders of consciousness” without being unfair to the skills that are left?

This work aims to discuss disorders of consciousness, using both philosophy of mind and practical knowledge in order to find a way to fairly treat and address patients that have not yet been diagnosed. To do this, this work will discuss disorders of consciousness, how they are diagnosed today and the main problems in clinical practice and observation.

Then, it will investigate any external markers that could be strong indications of consciousness, and from there speculate what kind of thing consciousness is and what theories are being discussed to try to unveil consciousness itself.

Finally, the conclusions will lead this work to discuss what we can do with the information that is available. In the final chapter, it will speculate whether there are any mathematical decision-making theories that could lead us to a final decision – the answer will be that this is not the best way to make decisions about disorders of consciousness, and therefore we must find a way to bring about ethical theories that will take into consideration a case-to-case scenario, respecting patient (or family) history and points of view.

Keywords: consciousness, agency, disorders of consciousness, philosophy of mind, bioethics

RESUMO

O que acontece quando a consciência falha em atingir seu nível pleno e supostamente “mais alto”? Quanta consciência é preciso ter para ser reconhecido como “consciente”? Poderíamos dizer que, se alguém não atende a um padrão claro de atividade consciente, não podemos considerá-lo um ser consciente? O que é suficiente para afirmar que alguém é consciente e como abordar os chamados “distúrbios da consciência” sem ser injustos com as habilidades que sobraram?

Este trabalho tem o objetivo de discutir os transtornos de consciência, usando tanto a filosofia da mente quanto o conhecimento prático, a fim de encontrar uma maneira de tratar e tratar com justiça os pacientes que ainda não cumpriram seu diagnóstico. Para tanto, este trabalho discutirá o que são distúrbios da consciência, como são diagnosticados hoje e quais são os principais problemas da prática clínica e da observação.

Então, este trabalho tentará investigar se existem quaisquer marcadores externos que possam ser fortes marcadores da consciência - e a partir daí, tentar especular o que é a consciência e quais teorias estão sendo discutidas para tentar desvelar a própria consciência.

As conclusões levarão este trabalho a discutir, por fim, o que podemos fazer com as informações de que dispomos. No capítulo final, tentará especular se existe alguma teoria matemática de tomada de decisão que possa nos levar a uma decisão final - a resposta será que esta não é a melhor maneira de tomar decisões sobre distúrbios de consciência, portanto, nós deve, então, encontrar uma forma de desenvolver teorias éticas que levem em consideração um cenário caso a caso, respeitando a história e os pontos de vista do paciente (ou família).

Palavras-chave: Consciência. Mente. Desordens da consciência. Filosofia da mente. Bioética.

TABLE OF CONTENTS

INTRODUCTION

> Perhaps no aspect of mind is more familiar or more puzzling than consciousness and our conscious experience of the self and the world. The problem of consciousness is arguably the central issue in current theorizing about the mind. Despite the lack of any agreed upon theory of consciousness, there is a widespread, if less than universal, consensus that an adequate account of the mind requires a clear understanding of it and its place in nature. We need to understand both what consciousness is and how it relates to other, nonconscious, aspects of reality. (Van Gulick, 2018)

One of the hard problems for philosophy is consciousness. Consciousness is usually described by philosophers and neuroscientists as a feature of “normal” human brains – the fact of being aware of yourself and your environment, and using it for action and decisions. But it seems so much more than that. Some neuroscientists say that consciousness is an “experiential” feature of individual brains. Nevertheless, this phenomenon does not seem to be explainable, at first glance, by just looking at the biological brain. Possibly, it is a category mistake to say that brains have conscious experiences – we will discuss this in chapter three. After all, consciousness is not an objective feature of human brains, but a subjective first-person experience of individuals (Nozick, 2001). How and why an individual’s biological brain transforms into a mind trigger many questions.

One of the puzzling problems of consciousness is that it is a private experience (Block, 1995). There is no way to access the content of other’s mental states – not objectively, at least. Agency is nevertheless an aspect of human behavior, something that can be objectively described. For that reason, some philosophers think that certain features of minds, such as agency, are almost fundamental keys for detecting consciousness fairly. But what happens when our capacity to express agency is not “strong” enough to present evidence that our agency is there? Does that mean we are not conscious at all?

What does it mean to be conscious, after all? Is it an “all or nothing” skill, or is there something that is like being “half”, “almost” or “quasi” conscious?

Neurologists often claim that diagnosing disorders of consciousness is a hard job – even with all the exams available, time and clinical observations are often used in diagnosing a disorder – and sometimes it can fail (Farah, 2013). The fact that we still have much to learn from our brains and minds and how they work together can be decisive for our often misleading ways to diagnose and deal with disorders of consciousness.

Thus, there are many contemporary cases that prove that there is much yet to be accomplished in this field. Diagnosing disorders of consciousness often takes several months, and sometimes what was believed to be one state turns out to be another. There have been several reports of what medical staff believed to be a comatose state turning out to be locked-in syndrome – states that are far from anywhere close in terms of what we believe to be consciousness.

Philosophy often plays an important role in discussing both minds and how to act upon medical ethics and discussing what forms consciousness, and how to perceive consciousness is a field that is often shared with many other areas of expertise. If understanding consciousness is something greater than the brains – physical organs – and needs the understanding of mental states and "what it is like" to be someone, philosophy might be decisive for understanding how minds possess consciousness.

In order to truly understand the challenges involved in this work we need to dive into the very early aspects of medical neurology and then discuss how this fits together with the best philosophical theories of mind presently available.

In this work, we will not take consciousness as being an "all or nothing" thing or try to reduce it to a "material" thing tied to its functions. Reductionism often tends to worsen problems when it comes to consciousness, and more often accomplishes nothing when trying to make explanations about something that we do not yet understand. Consciousness is a complex matter – for not only the biological but also the social sciences. This is a reason why neuroscience is today understood as a multidisciplinary field. Nevertheless, there are many ways to study and approach the field of the brain and mind, although all of them still do not solve the problem in all its complexities – hence,

there is no good way (up until now, at least) to reduce the problem of consciousness to only one discipline without being unfair to it.

But reducing this problem, as Nagel suggests (1974), is a logical temptation. We often think of things we cannot understand through analogies. Nagel says that this is unlikely to help us in the search of any useful answer on the matter of the mind-body and consciousness problem, and only lead philosophers to accept and discuss mental theories that are implausible and too reduced, and that do not help at all in the practical fields.

> Consciousness is what makes the mind-body problem really intractable. Perhaps that is why current discussions of the problem give it little attention or get it obviously wrong […] Every reductionist has his favorite analogy from modern science. It is most unlikely that any of these unrelated examples of successful reduction will shed light on the relation of mind to brain. But philosophers share the general human weakness for explanations of what is incomprehensible in terms suited for what is familiar and well understood, though entirely different. This has led to the acceptance of implausible accounts of the mental largely because they would permit familiar kinds of reduction. (Nagel,1974)

Neuroscientists have been trying to find a way to point out where consciousness in the brain and its mechanism are, but it is a phenomenon that is far from being understood mostly due to its subjective nature. One practical and honest way scientists have been trying to make sense of consciousness is to “corelate” mental states with brain regions. This, of course, can be somewhat studied in patients with brain damage, and there are some good approximations.

Nevertheless, correlations, although particularly helpful in beginning to understand how consciousness works, are only the beginning of the answer. If we consider that correlating a brain region with a function is doable using imagery exams, we may imagine how the physical brain is divided and how damaging one area might increase the risk of a certain symptom – for example, how damaging the frontal lobe might cause disorders of the emotions and personality. We *know* this by correlation, so when a person damages this area we often expect a certain cluster of symptoms associated with the damage. Nevertheless, this correlation does not explain how the information itself is processed and

what the experience might be when it comes to the result of this disorder on the emotions or personality from a first-person point of view – there are tests that might help us understand this, but they are all from a third-person point of view, and might not be reliable depending on the extent of the damage (we will come back to this in chapter two).

In this work, I will try to understand the best methods we have to indicate consciousness – if there are any – and how to use the results in a way that are trustworthy and fair. In doing so, we will begin to dispatch a philosophy of mind and enter into a bioethical discussion. Nevertheless, in the complex matter of consciousness, it does not seem plausible to consider this matter only within the limits of philosophy of mind – after all, the problem of consciousness is not a mere conceptual problem, but an empirical one. As an empirical problem, consciousness is also a problem that truly matters for ethics. So, in this work, I will suggest binding issues of philosophy of mind and bioethics in order to develop a way to address one aspect of this complex matter – that is, the problem of consciousness viewed from the point of view of studying severe disorders of consciousness. In order to do that, I will discuss and combine what we know from philosophy of mind with an approach of bioethics in a way that bioethics "modulates" the scientific knowledge enough for us not to end up with a course of action that seems too harsh to apply to the real world.

The role of philosophy here is very important, and, although there aren't many works of philosophy of mind that consider ethical approaches, it seems impossible to reach a result that could be fair and usable in healthcare without considering both fields. This will be the aim of this work. We do not have all the answers to disorders of consciousness and there are still problems that need to be discussed, even if we do not answer all the metaphysical questions that this work raises.

To reach my goal we will begin by discussing disorders of consciousness and how and why they happen in the first place. By doing this, I plan to point out why it is not an easy diagnosis and might not be so clear at first glance. All these issues might affect the prognosis of a disorder because in the practical field prognosis depends actively on positive actions. Pointing out those problems will lead me to one of the main questions of

this work – how are we to solve the problem of disorders of consciousness in a fair way in the practical field?

I will then try to understand how consciousness is viewed by philosophers in the field of philosophy of mind and discuss the progresses made by some like Ned Block and Tyler Burge. Understanding these theories will not only help us to understand the progresses being made but also give us the opportunity to relate what might be useful in the practical field and what is not.

After that, I will explore the possibility of external tests being good or bad markers, using one that is widely accepted as a marker of consciousness – agency. If it is the case that agency is a good marker of consciousness then we might have a fair way to determine how good a prognosis might be and what is "fair" for investing in the patient. However, I will find in the chapter that agency, although a marker of consciousness, is not a very good one, so the problem remains.

In chapter three I will then turn to some more adventurous theories about what consciousness might be and why considering it might help us with the practical fields and the understanding of disorders of consciousness. I believe I may have found some possibilities that can shine a light on the answer, and David Chalmers and Daniel Dennett will be allies in discussing my proposition.

Since there is a problem that I won't be able to overcome in this work – that is, proving that my theory might be correct, and moreover how to test for consciousness in a final-defining way in the practical fields – I will take chapter four to discuss what can be done with what we have so far, and try to be fair to the problem without causing ethical issues in the practical field.

I believe I have reached many of my goals – I have made a new proposition for what I think consciousness might be, and tried to solve the problem with what we have so far without being unfair to the practical field of the diagnosis and prognosis of disorders of consciousness. It is important to state that this is a work of its time – much will probably be unveiled in the years and decades to come. Furthermore, there is still much to be made in this field and on the subjects debated here in this work, and my contribution is

still very small. But every step is important and has to be made. Only that way will we one day, maybe, solve the issue for good.

CHAPTER ONE

What are disorders of consciousness?

> There is an irreducible philosophical limitation in knowing for certain whether any other being possesses a conscious life. Consciousness is a multifaceted subjective first-person experience and clinical evaluation is limited to evaluating patients' responsiveness to the environment. (Laureys, 2004)

There are several diagnoses when it comes to diagnosing severe disorders of consciousness – some are more similar, some further apart from each other. Neurology (and neurosciences) has been trying for many years now to best describe and perceive the signs and symptoms related to it, and the best way to tell whether someone does or does not fall into this category of disorders of consciousness. There has been some success in doing so, but there is still much left to do (Lange, 2014. Farah, 2013). Diagnosing disorders of consciousness is not an easy job – after all, most neurologists (Farah, 2013) would agree that discussing a problem that is yet not fully understood defies the very limits of medicine, to say the least. Are we right when we claim one state over a diagnosis? Is this fair to the allocation of resources for a decent prognosis?

Refining and understanding what disorders of consciousness are and *what it is like to have it* may have more to do with the field of philosophy of mind and bioethics than the bare medical methods themselves. To begin with, let us make it clear that all disorders of consciousness affect, at some level, the person's ability to *show personality* and ***experience*** *the world*. Some are more severe than others, but still it seems clear that these diagnoses share those two things in common. Due to this very particular nature, many questions about "how much of a person is left" are still being asked, along with whether they have experiences or not. We shall discuss this in this work.

The first thing to do here is to make it clear from the start that many things can lead a person to a disorder of consciousness. To name only a few: strokes and vascular accidents, mental illnesses, severe dementias, neonatal anoxia, infectious diseases, and

traumatic events. It is not the cause that binds them to this category of disorder of consciousness (DC), but how it is reflected in the brain (Farah, 2013. Lange, 2014).

Let us imagine a case to illustrate this. A person suffers a car accident. In this accident, she hits her head hard enough to put her in a coma. A part of her brain is severely affected by it and, in some ways, the brain is no longer doing what it is supposed to do. After a couple of months, this person cannot walk, talk, or express herself in any way – as a matter of fact, she is diagnosed with being in a vegetative state months after the accident.

The vegetative state, in this example, is not the “first occurrence” itself, but has been caused by it. We are right to think that the car accident did cause head trauma that eventually led to the vegetative state, though other circumstances can lead to this same severe mental disability. Another example could be an infant born with a severe perinatal anoxia and is also diagnosed, after some time, as being in a vegetative state. The cause here, although there are some distinctions worth going through, is not the DC itself but nearly the cause of it. This is just an example of how different “accidents” or “injuries” can lead people to the “same” state.

The distinctions can be named in a temporal distance – some illnesses are very long and demanding and take a long period of time to “put” someone in a DC, such as Alzheimer’s disease, for example, but some are more similar to an event – like a car accident or a stroke. The outcome, however, for all these situations tends to be better when the time of response is fast (Lange, 2014). Recent studies suggest that receiving treatment in the first few hours after the incident may increase the change of a significant recovery – Hospital Sírio-Libanês (2018) suggests that if a patient comes into an emergency room (ER) with symptoms of a stroke he should receive medical assistance within five to ten minutes.

If you find yourself in an emergency room and need to determine how “conscious” or “aware” someone is, the first test used to describe the amount of damage is the Glasgow Scale.

SIGN Guideline No. 46.[15]			
Feature	Scale responses		Score notation
Eye opening	Spontaneous		4
	To speech		3
	To pain		2
	None		1
Verbal response	Orientated		5
	Confused conversation		4
	Words (inappropriate)		3
	Sounds (incomprehensible)		2
	None		1
Best motor response	Obey commands		6
	Localise pain		5
	Flexion	Normal	4
		Abnormal	3
	Extend		2
	None		1
Total Coma 'Score'			3/15–15/15

*Barlow, Philip. “A practical review of the Glasgow Coma Scale and Score.” *The Surgeon: Journal of the Royal Colleges of Surgeons of Edinburgh and Ireland* 10, 2 (2012): 114–9.

The Glasgow Coma Scale[1] is a simple and very effective tool to stratify “how much” a person can respond and show, on some level, facial, verbal and motor responses to some stimulus, such as eye contact, sounds and communication and pain. Although this is an indirect test (it depends on how much the person can respond to it and is definitely not a direct way to evaluate consciousness), this is commonly used in emergency rooms and has proven to be very handy as a first response to traumatic/decreased consciousness events.

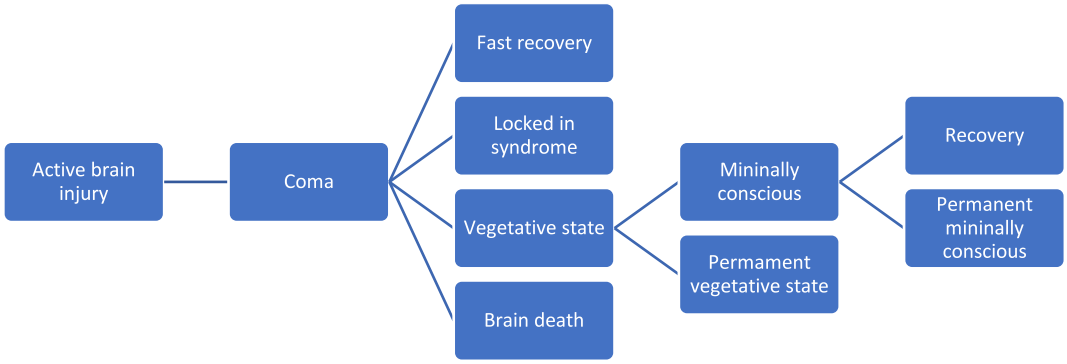
The Glasgow Scale is a numerical scale that ranges from 3 to 15, and the lower a person is on the scale the further into a “coma” they are, in theory. A coma is a state in which a person is unconscious due to a brain trauma or injury (the causes might vary from a physical trauma to intoxication, a stroke, infections, and so on) – comas might also be induced to lower the brain’s activity in order to prevent further brain damage (Lange, 2014). Comas might be temporary or develop into a major disorder of consciousness – such as vegetative states or minimally conscious states (Lange, 2014. HOSPITAL SÍRIO-LIBANÊS, 2018). We would hope that “healthy” people with no brain

[1] Notice that there is a specific Glasgow Coma Scale for infants and children due to their lacking the capacity to fulfill some of the categories of the ordinary scale, such as “oriented verbal response”. See JAMES, K. C.; PINA-GARZA J. E. **Fenichel’s Clinical Pediatric Neurology: A Signs and Symptoms Approach.** 8th edition. Elsevier. 2019.

injuries would get a “15” score, but as a subject progressively loses her ability to show conscious motor responses she will score less points. A person who is brain dead, for example, would be a clear 3. But a person in a vegetative state, might be somewhere between a 3 and a 7. A DC is a late occurrence that starts in other illnesses.

DCs have a history. In my example above, it started with a car accident that caused a head trauma that led to a coma and eventually a vegetative state (VS).

A DC has causes and is usually a result of something that is happening within – it **can be the result of an active brain injury**. As I said above, it is not the active brain injury itself that is the disorder of consciousness, but the results that follow from it. Different sorts of DC can be caused by a same general type of injury.



But how can we recognize such states? How can we tell them apart?

When it comes to making a diagnosis, physicians use different diagnostic tools to reach a conclusion. Even though today we have more accurate methods than in the past, diagnosis is still a complex matter, and we will look at why that is. I am not suggesting here that it is impossible to make an accurate diagnosis, only that it is very hard, even for a very experienced neurologist, to tell some cases apart.

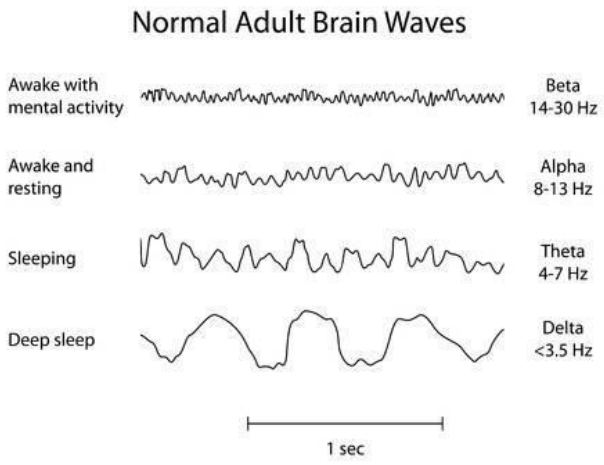
In the ER set, the first thing to do in acute cases is usually run some quick tests to discard things that are “easier” to treat – such as low blood sugar or infectious diseases (Lange, 2014). Also, having a full individual clinical history and family history might be particularly helpful in cases that involve hereditary conditions and dementias. But when it

comes to investigating chronic brain damages, evaluating how and how much the brain is “working” might come down to two functional exams: electroencephalogram and fMRI.

But when it comes to investigating chronic brain damages, evaluating how and how much the brain is “working” might come down to two functional exams: electroencephalogram and fMRI.

An advantage of the electroencephalogram is that it is a completely non-invasive method. The exam is carried out by placing electrodes on the patient’s head (sometimes with the help of a “gel” to fix them) and the machine measures the wave of electrical activity in every pulse fired in some region of the brain. It is an exam that measures electrical brain activity – pulses that take information from one neuron to another. Although the electroencephalogram is not a good exam for measuring the “quality” or content of the information being carried, it is a very good way to measure if there are any pulses at all, and if they are a match for normal activity or not (Farah, 2013). Most activity falls into a frequency range of 1 to 20 Hz (Lange, 2014).

It is important to note that brain electrical activity in a normal person varies according to their state while taking the exam – for example, if the patient is awake or asleep. Depending on what the doctor wants to diagnose, he can ask for different types of electroencephalograms while the patient is awake or sleeping, or a mapping electroencephalogram which associates the electrical activity with the area responsible for the pulse with the help of a computer.



(a)

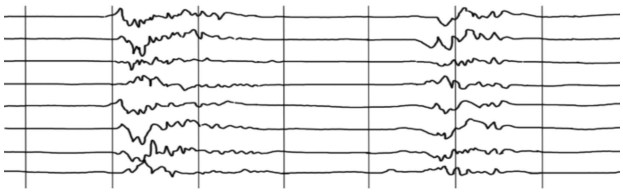
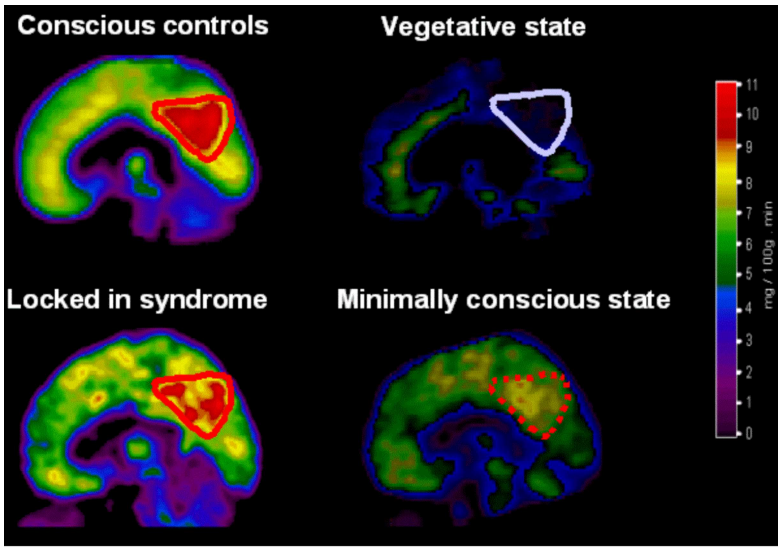


Figure (a) above is an electroencephalogram of a patient in a coma. Even if you are not a specialist you can probably see that there is something odd about the traces – the waves are widely spaced and do not seem to have a “normal” rhythm. What the waves stand for or mean depends on the state the patient is in (for example, awake or asleep) and their frequency.

Another very good exam, and much more “accurate” for finding *anatomic and functional* disorders, is Functional Magnetic Resonance Imaging (fMRI) which measures brain activity by mapping blood flow in the brain. The idea is very simple – if an area has blood flow or if the blood flow increases while doing some activity, we can conclude that that part of the brain is being used. fMRI can be a “more invasive” technique due to the use of a blood-oxygen-level dependent contrast. The result is a “picture” of the brain that is usually very colorful – a color-coded image that tell us what part of the brain is in use and how strong the diffusion of oxygenated blood is.

(b)



We can see some important distinctions in figure (b) above. In the conscious controls, for example, we can see very vivid and strong colors showing off the high level of oxygenated blood flowing through the brain. The brighter the color, the more oxygenated the blood, and therefore, presumably, the greater the "use" of the brain. Highlighted areas often tell us that this part of the brain is being used and, since we know what area is responsible for what, we can prove that some activity is taking place (Lange, 2014). For example, if the patient is placed in the fMRI and told to "speak," and the "speech" region of the brain lights up, we ought to think that everything is going well. The individual's brain is exhibiting neural activity to a proper stimulus, a neural activity that we assume is also a conscious activity.

Conscious controls, however, also tell us what not to expect in "normal" functioning brains. If a person is placed in the fMRI and told to "speak" or "move," and the part of the brain responsible for it does not show any activity, then we can conclude that something is wrong. If the colors are not as bright or if there is no color-coding at all we can assume that the neural function is compromised. Again, fMRI – in figure (b) it is also not a good way to show the "quality" of information going on inside the brain – is only a sign that oxygenated blood is flowing normally in the brain tissue.

One problem with diagnosing disorders of consciousness is that sometimes one test or exam is not enough and does not answer the quest for consciousness – many diagnoses are deeply dependent on time and clinical observations, and are therefore subject to interpretation. The healthcare team needs to make multiple clinical observations for a considerable amount of time (usually several months), using multiple tests and exams, which can be both expensive and time consuming (Farah, 2013).

The way we deal with determining disorders of consciousness is often misleading due to our limited capacity to fairly corelate the signs and symptoms with each single diagnosis. This happens because symptoms can be shared with multiple diagnoses, and the signs are often not clear enough for a fair diagnosis to be made (Farah, 2013). This does not mean that clinical diagnosing is not good (considering that we have a lot of difficulties at the present time), but it nevertheless shows us that we do not have an

accurate compass when it comes to the practical search for consciousness and its disorders.

Even if you are an experienced neurologist, the fact that symptoms can be shared among multiple diagnoses is a problem in itself. How to tell diagnoses apart if they can sometimes look the same? The way we choose to deal with this – with the fact that we still do not have a clear way to tell certain diagnoses apart – affects decision making in the practical fields, and consequently represents a direct consequence in the prognosis and investment of the patient themselves.

Because a prognosis is also deeply dependent on what we do in the early stages, mistaking one diagnosis or believing a patient has “a-state” instead of “b-state” can be a serious assumption with devastating consequences. The amount of resources we are willing to invest depends on the amount of certainty we have in the belief that a certain diagnosis can or cannot have a good prognosis, and early detection is a real problem here.

From what we have seen so far, fMRI suggests that a person with a “normal” fMRI test exhibits “normal” (as expected) states of consciousness – although the test itself only shows that the blood flow in the brain is normal, and nothing more. This result somehow implies that the normal blood flow in the brain corresponds to a “normal” state of consciousness. This leaves many questions unanswered – like what consciousness actually is, and what we understand and accept as consciousness, for example.

We will try to investigate these subjects, but to do so we might first need to take a look at some theories that try to explain what consciousness is – the progress in understanding what type of thing consciousness is and how it is composed, which might help us understand where we should look for an answer to the problem of treating disorders of consciousness fairly. We will begin by discussing the work of Ned Block, who has a fit first description for introducing the concepts of access and phenomenal consciousness. Block (1995) has made some progress in this field, and his definitions of access and phenomenal consciousness are still widely accepted (Burge, 2007), and we will discuss the problem from there, taking a few steps forward and away from Block’s work, and also bringing some new discussions in philosophy of mind into the picture.

Is all consciousness the same?

In this section we will discuss *Phenomenal Consciousness* and *Access Consciousness* and their relation to disorders of consciousness.

These two concepts were first described by the philosopher Ned Block in a famous paper published in 1995, and are largely used and discussed in philosophy of mind due to their progress in discussing consciousness functions. Although we will see in this section that Block's theory has a few problems, it is fair to say that his work is an important step in the understating of consciousness and its role, and how consciousness is composed. Block understands consciousness as divided into two models: *phenomenal consciousness* and *access consciousness.*

First, this is Block's definition for access consciousness (AC):

> A state is access conscious (A-conscious) if, in virtue of one's having the state, a representation of its content is (1) inferentially promiscuous, that is, poised for use as a premise in reasoning, (2) poised for rational control of action, and (3) poised for rational control of speech. (Block, 1995)

This is his definition of phenomenal consciousness (PC):

> P-consciousness properties are experiential ones. P-conscious states are experiential, that is, a state is P-conscious if it has experiential properties. The totality of the experiential properties of a state are "what it is like" to have it. Moving from synonyms to examples, we have P-conscious states when we see, hear, smell, taste, and have pains. P-conscious properties include the experiential properties of sensations, feelings, and perceptions, but I would also include thoughts, desires, and emotions. (Block, 1995)

Philosophers of mind are inclined to think that AC is a particularly important, if not essential feature of consciousness. Many argue that AC is all that matters for proving consciousness (Block, 1995. Boly, 2007). Undoubtedly, AC plays an important role in

what we perceive in another as “proof” of consciousness being present, but is, as Block argues, only strong because it is observable (1995). I partially agree with those who say that someone is only conscious if they have AC. Nevertheless, in this section I intend to present some objections and criticisms. It seems important, or it should matter, the extent to which we can expose our inner state. In this section I will argue that AC actually does play an important role in our lives and society, but it does not seem to be the case that all that matters for consciousness is AC, or that AC is a sufficient feature for saying that one is conscious or not.

Nonetheless, for Block (2004. 1995), AC seems to be a sufficient condition for stating the level of consciousness of a person. It is hard to argue against this, but let us see. For Block, if you have AC, we are authorized to say that you are conscious, that is that you have conscious states of mind. An ability such as AC certainly plays an important part in our lives. Asserting that we have this kind of mental ability can explain much of our human behavior. Hence, it is plausible that we display an ability such as AC, and since it seems prima facie plausible that PC is an ability distinct from AC, Block’s view seems to be at least conceptually reasonable. This view is actually very persuasive within philosophical circles; nevertheless, it seems not to be in agreement with what is observed in practical fields, such as the clinical medicine, which I intend to show.

But what is the difference between these two very different conscious states? When we imagine what conscious states are like we often think about how we feel toward being conscious (Nagel, 1974). We think of ourselves as having access to our experiences in the most complex way. In this case, this means having “access” to an experience that can be understood as a state in which we are aware of the experience itself (“what it is like to have it”). To be “poised” must be, as Ned Block claims (1995), the availability of accessing the experience itself. We do not think only if there is “something that is like to be us” – we consider, at first, access itself, or, in other words, if we have “experiences” that are available from our own point of view.

PC is a state in which there something that is like having it – raw experiences. If someone has something that is like being them, they have PC.

> But no matter how the form may vary, the fact that an organism has conscious experience at all means, basically, that there is something it is like to be that organism. There may be further implications about the form of the experience; there may even (though I doubt it) be implications about the behavior of the organism. But fundamentally an organism has conscious mental states if and only if there is something that it is to be that organism – something it is like for the organism. (Nagel, 1974)

So AC seems to be, at first glance, a consequence of PC. These states can be described as follows:

> *A state is **phenomenal** conscious if and only if there is something that is "what is it like to be" in this state; a state is **access** conscious if it is an available state for phenomenal mental processes or states involving reasoning or controlling of action or speech.*

Block argues for a strong interaction of AC and PC, but in his approach AC plays a more fundamental role than PC in the functional map of the human mind. Although they interact, and although AC, for Block (1995), is an "essential" feature of consciousness, PC seems to have a vivid role in how much we experience the world – or, in other words, what it is like to be ourselves.

Philosophers often take consciousness as being an "all or nothing" predicate, especially in the context of Block's PC/AC approach. PC is a state of mind of which we are very familiar. It is an intimate state, that correlates with how (and not how much) we experience the world daily when we are awake. PC seems to be metaphysically primitive compared to AC, and AC seems in some way to be a necessary expression or manifestation of PC – but this isn't always true. If it were, whatever AC states we were in would be necessarily accompanied by some cluster of PC states, and vice versa. In fact, disorders of consciousness are evidence that this cannot be true.

Besides that, the alleged difference between these two states is not sufficiently clear cut. The fact that PC does not imply availability for "what it is like" to have an experience does not rule out the fact that it has something that it is like having one. The fact that PC cannot be perceived (that we cannot have any public or objective evidence in its favor) cannot be mistaken for the lack of **any** consciousness at all. If we agree that

there are some states in which it has something that is like to be them, and if we value the fact that “there is something that it is like having it”, then PC would be the most important for us.

One decisive case is *locked-in syndrome*. This is a pathology where AC seems to be somehow compromised, but PC is supposedly fully functioning. If it is true that PC is qualitative, then it must also be true that there is something that is like having only PC.

For Block (2004), AC also has an important value, but it does not seem to be the case that it can display it alone – that is, following Block’s approach. AC seems to be supported by PC, and in this case PC is what really matters. Of course, it might be very difficult to argue “how good” the pure PC experience would be without any AC (we will come back to this point later), but it seems impossible to consider that there is such a thing as a lack off or a default AC erasing the possibility of PC being there in the brain.

Another important question here is whether PC can have multiple ways of showing itself. I believe that this is the case. Let us imagine this:

A minimally conscious case: Jade had severe anoxia during her birth. She had a brain injury that compromised many important brain functions. She cannot talk, but she can hear and even show emotions – she smiles while watching TV and cries if she feels pain. Although Jade is already ten years old, she cannot do what most ten-year-old girls do. She spends most of days in bed, since she cannot walk, being cared for. She experiences some joys and some pains, but we can, for sure, tell by her exams that her brain is not fully “normal”. As a matter of fact, Jade was diagnosed as being minimally conscious.

Jade’s case shows that she has some experiences that correlate with PC, but does not have them all. She is limited of course, but does somehow have some PC and even some poised experiences that would be classed as AC. Her case shows us how PC may not be as complete as we previous thought it ought to be. Jade, in her own way, does

experience the world “enough” to have something that is like being Jade. We would be okay with the position of:

There is something that is like being Jade.

We can argue about “how much” consciousness is left in Jade’s case, and that is another good discussion, but we certainly cannot perceive Jade as not being conscious at all. Jade does have some consciousness, but it does not seem to be the case that we can rule her out from having PC or even AC due to her “deficiency”. Jade’s experiences are different – rudimentary, some might say – but consciousness is definitely there.

Now let us look at a case that is further down the Glasgow Scale:

The severe vegetative state: Nelson has been in a vegetative state for a couple of years. He has a massive brain injury. He opens his eyes but does not focus on anything. He reacts to nothing other than pain. His fMRI shows little activity, and his physical exam shows some signs of decerebration. He does not make any noise, does not eat alone, and does not interact in any way.

In Nelson’s case, it is awfully hard to argue that there is much of a PC left, and AC is even harder to argue. Jade’s case, compared to this one, is much simpler. There is something that is like being Jade, even if Jade’s functions are compromised at many levels. Nelson, on the other hand, is something more severe. His fMRI and the signs of decerebration support that there is little or nothing left to prove that there is something that is like being Nelson. So, in this case, we would be correct to assume:

There is nothing that is like being Nelson.

Nelson's interactions – access consciousness – are close to zero, if not at zero. The only thing we can attribute to Nelson is pain. And this is a very hard discussion, because in Nelson's case it is very easy to associate pain as being a "reflex". We are not able to prove that Nelson does have PC when it comes to pain.

In the case of Nelson having phenomenal consciousness with pain, we ought to think that Nelson's life is a very bad one. In this case, we could assume:

There is nothing that is like being Nelson, other than pain.

Nagel would nevertheless say that "the more different from oneself the other experiencer is, the less success one can expect with this enterprise" (1974). In the sense that we cannot share first-person experiences, the quality of the experience itself is impossible to determine. We can only (according to Nagel) try and imagine a "sufficiently similar" (1974) scenario that would give us a clue to how (or try) to understand first-person experiences from the third-person viewpoint – this would be the best approximation.

So, the comparison that it is important to make here is this – it is true that some patients, while undergoing surgery, are in that moment in a state that is very similar to Nelson's. If they do have pains we ought to think that they will feel them, and therefore will not have PC that is related to pain in that moment. Nevertheless, Nelson's case is vastly different from the patient undergoing surgery. There is something that is like being the patient undergoing surgery – they just lack, in a part of their history, AC and PC. They will most likely recover and go on with their lives, while Nelson most likely will not. Temporary cases and definitive cases, such as the case of permanent vegetative states, cannot be compared here.

The important comparison here is whether "not feeling pain" is a good argument for not caring about pain if the person does not remember feeling it – like in surgery, for example – or if reflexes or raw body sensations are worth discussing in the realm of philosophy. Perhaps, if we care about "how well" our lives go (Parfit, 1984), then it might

be the case that even if we do not remember pain it would still matter. PC is, in that case, very decisive, but not all that matters. If it is the case that Nelson does not have PC when it comes to pain, we still ought to think that his life is bad – a life most likely not worth living. And if he does, then it seems to be even worse:

There is nothing that is like being Nelson = Nelson does not have PC, and therefore there is nothing that is like being him

but, if:

There is nothing that is like being Nelson, other than pain = Nelson might have some PC, and therefore Nelson's life is only pain.

The many faces of consciousness

I suggested in the previous section that there must be different faces to phenomenal consciousness, and now say the same for access consciousness. Perhaps one good word to use would be "levels," but not in the sense that there is a staircase in which we can classify people when it comes to phenomenal or access consciousness "skills," but that there can be skills that are there without being necessarily attached to others shared by the same class (phenomenal or access).

P-consciousness	**A-consciousness**
Hear, smell, taste, feel, sight	"Poised" reasoning
Pain	"Poised" rational action
Thoughts	"Poised" rational speech
Desires	
Emotions, sensations, perceptions	

Let us take PC first. It is not hard to imagine one – let's name them "skills" – existing without the other. You can imagine, for example, a blind person that has something that is like to be she. Consider a person who is deaf or who lacks the capacity to smell due to some pathology. This is a disability, but there is still something that is like being them. Thoughts, desires or emotions can be equally affected by depression, akinetic mutism, dementias or cognitive disabilities. It is not the case that a person who has depression, for example, does not have other skills or something that is like being them.

One important question is what it means to be "poised" to do something. If Block (1995) was referring to a second-level state, it must be dependent upon a first-level state. AC seems for Block to be dependent on PC – at least a part of it. If that is true it means that everyone who has access consciousness must possess at least some facet of PC. Now, it does not necessarily seem true that access consciousness needs to run on a very strong PC (Burge, 2007). If that were the case, minimal state consciousness would not exist. Think like this opens a new world of possibilities for what it means to be conscious.

If we think of consciousness as being necessarily a full and high-level skill only, it is true that we cannot think of minimal-state consciousness as being consciousness at all. But that does not seem to be right. If someone in a minimal state of consciousness can have at least minimal mental states poised to do something (maybe quasi-inferences, poor sketches of actions or perhaps global aphasic communication attempts), and those "poised states" come to PC states, then it must be true that they have some kind of AC, and that the subject is therefore experiencing something that is like being them – and this cannot be called a non-conscious state of mind.

We often think of consciousness as a very hard concept, but it is also true that when it comes to analyzing consciousness we attribute a conscious-like status to a lot of different "states." For example, we often think of infants as being conscious, even if we recognize that they do not experience something that is like being them in the same way a mature person would experience it. Recognizing children as conscious is not problematic for us, even though some may agree that infants and children are still building skills for "high order" consciousness, as Harry Frankfurt (1971) suggests when he says that children still are not "persons" but "wontons".

AC and PC are parts of what we believe composes some type of “full” consciousness – but are they consciousness proper? Is that enough to prove someone as being conscious? The best guess is no. AC and PC do help to clarify the discussion, but they find it difficult to say what goes along with them when it comes to practical fields. This might be because AC and PC are “features” or “characteristics” that we attribute to consciousness – parts of a whole. Nevertheless, they can still be hard to prove, in the sense that they can be hidden, difficult to observe and even more difficult to test.

I claimed that PC is severely mistaken as a very rudimentary and not-as-important skill. As a matter of fact, Block clearly favors AC over PC when it comes to degree of importance. This might be a little too soon for favor access so much – if AC being more important than PC were the case then we would not consider animal ethics as an important issue that would require our attention, and yet we do spend time (a lot of time) discussing it. It seems obvious that we do not think of animals as being “poised” in the way that Block describes, but most of us think of animal ethics as a subject that at least deserves some of our attention and fairness. If that is the case then it must be because most animals, at least the ones with spinal cords, must have something that is like being them. We seem to recognize that animals can suffer, feel joy or appreciate life itself – even if some of us consider them as being not entitled to AC at all.

We can say the same for toddlers and infants when it comes to AC, and even if we cannot understand how well a baby experiences things, our feeling is that their experiences still matter. Of course, there are some differences between animals, toddlers and disorders of consciousness, but these are just examples. There is a potentiality massive difference in an infant, for example, in the fact that they can later develop a full AC (or some of it), while most animals do not seem to have this capacity. My point here is not to discuss those differences in those subjects, but to point out the similarities to prove that the discussion of how good the experience is does not rule out the fact that there are still experiences going on, and this matters to us. If it does matter how good or successful our lives may be, then we ought to think that enhancing and valuing experience are the main goals of this work.

The hardest case here would be Nelson's – if there is nothing that is like being Nelson, then there is no experience that can correlate with Nelson's existence. In this case, Nelson has nothing like being him, and therefore we cannot even argue about "how good" Nelson's experiences are – they seem not to exist at all. Decortication and decerebration are perhaps much more vivid cases in the realm of bioethics then in that of the philosophy of mind – it is left for bioethics to advocate on what to do in such a case.

If it is the case that Nelson has nothing that it is like being him, other than pain, then pain is all Nelson has for experiences. If it is the case that Nelson's pain is only a reflex, it is still true that Nelson reacts to pain, at least physically. If he still does not feel it, it seems unpalatable to defend that Nelson's life has experiences that are good enough to justify it as something worth investing in. If we think of a patient in surgery who does not remember being in pain, we still refer to the pain as something that is intrinsically bad within itself. If pain has benefits, such as being in pain for an amount of time (let's say for surgery) and then recovering, then we accept pain as something that does not harm us – but still, not harming is different from something that does not matter. We avoid pain – sometimes even if there are benefits to it. Physicians often tend to postpone surgery, for example, if there are other ways to treat an illness – even if it sometimes takes longer.

Pain is something that really matters to us, and seems to matter even if we do not remember it. If someone feels pain and then cannot remember being in pain this is still a subject deeply related to ethics. Imagine this – you undergo a study about pain, and the researchers tell you that they are going to stab you in your arm with a needle. You will be in pain for some small amount of time. They also tell you that you should not be afraid because they are going to give you a drug afterward that is going to erase the memory of ever being in pain. You will have the experience of pain but not remember it. Would you do it? If you care about those few moments of pain, you would probably walk out.

Some might say this is irrational because you will not remember it – but still, we cannot advocate for this being a "strong" argument for causing or letting someone be in pain. From the point of view of PC, you will for a moment have something that is like being in pain. If that is the case, and if there is something that is like being in pain, then this matters.

If we are interested in how well our lives go, PC cannot be unseen – Thomas Nagel argues for PC having “something that is like to be that organism”. If that is the case, and if he is right, this matters.

Although we often believe that philosophy of mind is only a theoretical field with some approaches to the practical fields, the work being done in the realm of theories of mind and consciousness closely relates to practical fields, and there is much that can be helpful (and applicable) to the fields of practical knowledge. So, we still need to be very careful about advocating for something that can be potentially problematic for affairs that impact people in the practical field – pain is a huge thing, and should be taken very seriously, even only in theory.

Moving forward, Block’s exercise of understanding clarifies what to look for when it comes to markers of consciousness – which we will see in the next chapter. Nevertheless, understanding what AC is and is not and what PC is and is not does not answer how we can tell if someone is conscious or not.

Block might be right when it comes to the existence of AC and PC, but because of this it does not mean that finding AC and PC and pulling them apart in the practical fields are easy tasks. It follows that some important questions need to be asked based on Block’s work:

1. *Is there any type of PC without AC?*
2. *Is there any type of AC without PC?*
3. *What is the experience like for the “beings” that lack one of those?*
4. *Is a “being” conscious in some way if they lack some of these attributes?*
5. *How can we perceive others’ AC and PC?*
6. *If we cannot perceive them in some situations, does that mean that they do not exist?*

These six questions often prove to be harder to answer than philosophers had previously thought. Block argues in favor of AC existing without PC and vice versa. Nevertheless, this still does not help to clarify the quest of spotting them and deciding who is conscious and who is not. In addition, there is still the question of what the experience is like for the beings that lack one of them.

In order to prove that someone is conscious or not we need something "external" and "observable" – something that we can measure and test. We need markers.

Markers are, at this point of the work, something that can refer to the task of inferring someone as being conscious or not – a step that Block's work cannot achieve. A marker is not consciousness itself, it is evidence that indicates if consciousness is present or not, at best guess. There is some particularly good work being done in this field (Bayne, 2012. Klein, 2017).

It is good to clarify this by saying that markers are indirect ways for proving consciousness, but can also be a good helping hand in the practical field (we certainly use this in everyday life without noticing it).

What consciousness is and is not

This discussion goes back to Block's distinction between AC and PC. First, he suggests that access consciousness seems to be a development of phenomenal consciousness. I do not totally agree with that.

Block's contribution to this field, although valuable, is not quite correct all the way through. As Burge (2007) says in his book *The Foundations of Mind*, "access conscious states, and even events, need not themselves be phenomenally conscious" (p.383), and "although phenomenal qualities are individuated in terms of what it is like to feel or be conscious of them, one may have phenomenal states or events with phenomenal qualities that one is unconscious of. Thus, phenomenal qualities themselves do not guarantee phenomenal consciousness".

One of the problems of believing PC to be “experiential” is that “experience” does not seem to be only a phenomenal event – it can be **also** an access event. “What it is like to have it”, in this sense, needs some type of “sensation” that is too complex to be explained by PC only. In a sense “what it is like to have it” seems to depend on a type of AC – maybe not a complex one, but still some kind of “raw simple AC”.

We might be prone to believe that someone in a vegetative state “is able to have pains” and react to it physically – the heartrate might increase, for example, but they still won’t be able to “feel it enough to have something that is like having it”. Burge suggests here that the individual that I have just described lacks what he calls “rational-access consciousness” (RAC), meaning states that are, in this approach, intentional states.

The problem here is that phenomenal properties may or may not be “poised” by the individual. Like Burge (2007, p.389) suggests, “I distinguish what-it-is-likeness (phenomenality) from what it is concurrently like for the individual (phenomenal consciousness). In this sense, phenomenal consciousness is not enough for what we believe consciousness to be.” Individuals with phenomenal consciousness and even individuals that are, in a sense, “poised” may be unconscious.

> But phenomenal consciousness is fundamental to typing phenomenal properties, and phenomenal properties are fundamental to typing phenomenal mental states and events. The way a pain feels is essential or basic to what pain, and what a pain, is. The same is true with other sensations and feelings. I think nothing could be a pain, a token event of pain, and lack the what-it-is-likeness or characteristic feel or phenomenal properties that individuate pain. (Burge, 2007, p.389)

> First, to be rational-access conscious, a state or event must be poised for use in the central rational operations of an individual (animal or person). Second, rational-access consciousness must maintain at least a general connection to phenomenal consciousness in the individual. I do not, however, believe that to be rational-access conscious, a state or event must be phenomenally conscious. The connection between the two kinds of consciousness is loose, though phenomenal consciousness often seems to be a factor in the consciousness of rational-access conscious events (two types of consciousness). (Block, 1995)

PC is, of course, important to consciousness, and plays a role that cannot be reduced by AC. “What it is like to be” defines “what it is to have (a feeling)”. If I do not

know what it is like to be or to have a feeling, there is no access and no RAC, maybe just a raw PC. PC may comprise phenomenal properties. If you do not know what taste is, if you do not have phenomenal experiences with taste (ageusia), for example, you cannot know what it is like to sense it – and this can never be poised for use.

My point here is that AC is not only a development of PC, as one person can have instances of AC without PC, and vice versa. On the other hand, PC without AC can be “a raw PC” or a RAC. Those states, of one without the other, are disorders of consciousness – not a normal state, but they do exist and are totally possible to find in the practical fields.

Consciousness, in this sense, is composed of multiple aspects – not simply AC and PC themselves, but also those that may have many faces and aspects. Disorders of consciousness, and consciousness itself, are that complex.

CHAPTER TWO

Agency, intention (self-governance) and the role of consciousness

> The concept of agency is important for philosophers (who are interested in understanding what it means to be a self-determining being) and for cognitive scientists and psychologists (who seek to understand, for example, how some people can come to lack some of the attributes that we associate with fully realized autonomous agents, and how to prevent and treat such conditions). (Millican & Wooldridge, 2014, p.3).
>
> Agency is no less multifaceted than consciousness. The folk-psychological terms that we have for describing agency – "intentional agency", "goal-directed agency", "voluntary agency", "deliberative agency", and so on – are imprecise in various ways, and it is unclear which, if any of them, will find gainful employment within the scientific study of agency. (Bayne, 2012)

One of the questions for this work is the role of consciousness in agency and vice versa. If it is true that consciousness plays an important role in agency, then we should clarify what it means to be conscious in the context of agency. In this section, we will discuss this.

One of the problems of consciousness is that consciousness is a private experience and there is no good way to truly access the contents of private mental states. We often claim consciousness to be attached to some sense of being "an agent" – in other words an action, or with some intention towards acting in a rational, chosen way, or having responsibility for an action. At first glance, this seems to be a good way – and many philosophers would agree – to prove someone to be conscious. But is it decisive?

The first barrier prior to discussing agency is that the concept of agency is as diverse as that of consciousness itself. There are many "well-accepted" concepts of agency, none of which help philosophers with deciding what agency is at its core.

Agency can traditionally be understood as a "capacity to act", but not only for the sake of acting, but as having intentionality to act – consciously choosing to act (Schlosser, 2019). This is a useful tool for philosophers for recognizing agents in the world (in a very shallow way, in my opinion), and therefore imputing to the agent the responsibility of the

action itself, which can be understood as simply “an action that is initiated intentionally by the individual”.

Nevertheless, imputing so much on the agent becomes controversial when we go deeper into the discussion of agency – it is, in a second and more careful instance, a question of “how much” we are aware of what we do, and if we can determine what is at stake when we say that “someone” is responsible for “a-action” – specially, as in this work, dealing with disorders of consciousness.

This matter of “how much” one is responsible or aware of the action itself is not so agreed on when it comes to stablishing a theory. We certainly seem to be prone to attributing meaning to actions, and believing in the illusion of agency as a “controlled” subject in many situations – but we can be wrong. Some even claim that this puts “too much” pressure and responsibility on the agent, and is simply too much to be plausible.

Several arguments have been offered for or against this, but here I will focus on one that could help us in the discussion of consciousness and agency as a “marker or consciousness”. Let us focus first on Libet’s experiment (1993), which seems to conclude that unconscious events in the brain precede the consciousness of an intentional action. In this experiment, participants were commanded to initiate movement at their will.

This experiment has several points that are interesting for our subject. The participants were commanded to initiate action at their own will (which seems to be a problem already because, since they were commanded to do something, some could say this was not a choice, in a strict sense). Then, the participants initiated a bodily movement upon their will while having their EEG and EMG measurements recorded. Eventually, the EEG showed that there was neural activity before not only the muscle activities recorded by the EMG, but also what the participants recognized as the time of the unplanned will to act, and this seemed to predict the bodily movement before not only the bodily movement itself, but also the conscious will to move that part of the body, implying that an unconscious choice preceded the conscious will to choose, and therefore that there is no “free will” in the sense of a conscious intention unpreceded of nothing else, fully able to control one’s actions (Schlosser, 2019).

Seeing Libet's results, we may have some questions that help us scrutinize the results themselves, such as what is agency, what is intention and what is "command following"? Although Libet's experiment may be controversial for several reasons (which is not our focus here), it serves as an entry point for a proper understanding of the concepts of agency, intention and command following.

In order to work with the concept of agency, we do need to clarify what we are talking about, and what concept we are going to use. Agency can be used in both philosophy of mind as a "marker" in the search for consciousness and in ethics (and bioethics) to determine moral responsibility. None of these approaches are wrong, but they are as far apart as they can get. I will discuss aspects of consciousness and bioethics in a later chapter, but for now the topic will be *agency as a marker for consciousness*.

Even so, in philosophy of mind, there are a few concepts for agency, and many will be used in this chapter to enrich the discussion. Agency can have these perspectives or views:

• *The first-person view.* From this perspective, agents are purposeful originators of deliberate action, moved by conscious purposes.

• *Third-person view.* From this perspective, agents are entities whose behavior can be predicted and explained through the attribution to beliefs, desires and rational choice (Millican & Wooldridge, 2014).

Agency seems to be a good marker of consciousness at first glance, and here we can understand why philosophers might use this concept to determine if someone has consciousness in a normal state – an approach that, as Tim Bayne agrees (2012), is quite compelling.

This is important because, if we do find a marker – something that points out **for sure** that someone is conscious or not – the problem of diagnosing disorders of consciousness is solved, and the search for consciousness itself will also probably be closer to an end. This certainly seems to be an important step – the practical fields would benefit a lot from this – and the theory of consciousness (and mind) in philosophy would also allow for important steps forward.

A marker is something that serves to identify, predict or characterize something. We do have markers that we use in medicine and that are very good – such as the genetic markers of trisomy 18 and Down syndrome, for example – an amniocentesis can help us find this marker (an amniocentesis is a procedure where amniotic fluid is collected from the uterus of a pregnant woman and screened for fetal cells that provides information about the fetus itself, such as if she carries a trisomy or not). An amniocentesis can provide strong markers for those diseases. These are trustworthy exams and may diagnose the trisomy. If we find a genetic trace of the trisomy in the fluid then this is our strong marker. On the other hand, an ultrasound done for the exact same reason might measure the nasal bone, for example, and find that the measure is suggestive of a trisomy – the shorter the nasal bone, the stronger the marker – so if the nasal bone of the fetus is shorter but still close to normal size then this is a soft marker. Ultrasounds may also have distortions caused by the fluid around the fetus or the position of the fetus itself – so a marker found in an ultrasound is a soft marker if we compare it to a marker found in an amniocentesis.

But going back to the first and third-person views, I am going to invest a lot of time investigating the third-person view to evaluate agency as a good (or bad) marker, and since this is one of the main focuses of this work (pinning down how to spot consciousness), it is essential to evaluate what seems to be the most used theory in doing so (Bayne, 2012). We will try to set if agency is in fact a good marker of consciousness, and if it is decisive in the sense that it can be the only marker for pinning down whether someone is conscious or not.

The first-person view is also a concept that we could use alongside the third person – but since it can only be perceived by the person themselves it is not very useful for testing consciousness in the world. The focus is going to be on the third-person view for obvious reasons, such as the usefulness of the concept in helping to spot consciousness in the practical fields.

> An attractive place to look for such markers [of consciousness] is in the realm of agency. Consider the infant who reaches for a toy, the lioness who tracks a gazelle running across the savanna, or the climber who searches for a handhold in the

> cliff. In each case, it is tempting to assume that the creature in question is conscious of the perceptual features of their environment (the toy, the gazelle, the handhold) that guide their behavior. More generally, we might say that the exercise of intentional, goal-directed agency is a reliable guide to the presence of consciousness. To put it in a slogan, we might be tempted to treat agency as a marker of consciousness. (Bayne, 2012)

There are a few problems worth discussing here. One, if agency can truly be tied to consciousness, it would make it hard to prove an easy solution other than admitting that we are only conscious when we have agency. Nevertheless, if agency is not a good marker of consciousness, then a question is whether an agent might still leave room for conscious beings.

Philosophy often takes agency as being, in a way, what Block (1995) defines as being "poised to". There are many similarities between agency and what it is described as "poised", as it involves being intentionally motivated toward something or, in other words, to "mean" an action rationally (Block, 1995. Boly, 2007). Thus, the similarities only come this far.

Because it is so diverse in its meanings and because we cannot assume it to be simple, agency can be both a good marker of consciousness and, at least sometimes, a bad one. If we take the first-person view on agency above – that is, agents as *"purposeful originators of deliberate action, moved by conscious purpose"* – then we might assume that this "being" is conscious, as deliberative action stands for "self-governance" and reasoning, in which we can assume that it needs states that are truly invested in it (that is, in deliberative action), or in other words "poised" to action.

Nevertheless, if we take agents as "*entities whose behavior can be predicted and explained through the attribution to beliefs, desires and rational choice*", then there seems to be a lot more involved than just "deliberation". How can we be sure that there is some particular conscious mental state that is poised for some specified **voluntary action** to take place, especially if unmediated by any specific phenomenal state?

Consider this:

The blind case*: Carl has a partial cortex lesion. During a medical appointment he is put in front of some obstacles and asked what he is seeing in front of him. Carl says he sees nothing. The doctor asks him to come and sit by his desk to talk, but if Carl only moves forward he will stumble into the obstacles placed there. Carl moves away from the obstacles and makes his way to the desk without any incidents. Carl did not see or know what was right in front of him, but was somehow able to choose the best way.*

Blind patients are truly a wonderful interrogation for the status of consciousness – after all, are they conscious of their surroundings or not? And, more importantly, how can they access the information and deliberate if they are not "aware" of the process of decision making that we assume as a necessary condition for rational action? Conscious agency finds a big enemy in blindness:

> […] even when automatic actions are executed unconsciously, the stimuli that trigger them are typically conscious. One might not be conscious of initiating, guiding, or completing the action, but one will usually be aware of the environmental feature to which one is automatically responding. Think of the infamous long-distance truck driver who navigates "on autopilot". The driver might not be aware of making adjustments to the steering, but she will typically be aware of the features of her environment (the road, the stop signs, the traffic lights, and so on) that motivate such adjustments. (Bayne, 2013)

It seems that for whoever claims that agency is a good marker of consciousness, in the blindness case there should be enough information to infer that Carl deploys agency, and is therefore conscious. The trick here lies underneath all this – since we do not access Carl's mental states, we are fooled about what we observe. From a first-person point of view, for Carl:

There is nothing that is like choosing an obstacle-free way.

But, for the third-person point of view:

There seems to be something that is like choosing an obstacle-free way.

Agency, in the case of blind patients, is a two-way thing – they are both somehow "conscious" and "not conscious". It seems that Carl does not have mental states "poised" to action in the sense that he has taken a conscious voluntary decision, but since Carl is not certainly being "guided" by alien forces, what else could lead Carl other than he himself?

Of course, in the core sense there is something that is like being Carl – Carl does have PC, and certainly some type or level of AC. Blind patients usually have an impairment of the visual input-output cortex due to lesions in their striate cortex, also known as the primary visual cortex or V1, being able to respond to visual stimuli that they do not consciously see (Berkowitz, 2017). One question here is if volition or "a desire to a certain will to be his will", or in other words conscious voluntary decision making, is the only good marker of agency, and therefore of consciousness. If it is true that we can somehow say that Carl is conscious, then we need to ask ourselves what kind of consciousness Carl actually has:

> This brings us back to the issue of automaticity. In a very interesting discussion of this case, Levy (2008) takes issue with the claim that volition is evidence of consciousness. Levy points out that there is a large literature within both clinical and social psychology which suggests that action can occur independently of consciousness. For example, giving subjects stimuli that prime for thoughts of old age leads them to walk more slowly than control subjects. (Bayne, 2013)

This is also one of the main points of Harry Frankfurt (1971). Frankfurt classifies desires into two categories: first-order desires (and volitions) and second-order desires (and volitions). First-order desires are "core" and "rudimentary" sensations like being hungry and looking for something to eat. It is an action that is not mindful and does not

need a lot of effort in the sense of “want”. This primary state is, according to Frankfurt, shared with other animals like cats and dogs.

As for second-order desires, they are more complex – we can state they are what we consider to be “high-order” desires. For Frankfurt, “persons” are distinct entities that have also second-order desires besides the first-order ones:

> Someone has a desire of second order either when he wants simply to have a certain desire or when he wants a certain desire to be his will. In situations of the latter kind, I shall call his second-order desires “second-order volitions” or “volitions of second order”. (Frankfurt, 1971)

Can we classify the Frankfurtian second-order volitions as AC states? These states, being *wanting a certain desire to be one’s own first-order will*, seem to have a rational status – one that marks the individual as a *conscious agent*. For Frankfurt, this is an essential feature of being a *person* and not a *wantom*.

Frankfurt seems to be excluding many people from what he calls “persons” – toddlers and infants, for example. Although this might seem controversial, it is also a smart move since, in ethics, agency is often close to moral responsibility, and we actually do not think that toddlers and infants are fully responsible agents.

But, if we take Frankfurt’s view along with Block’s, we can go further as maybe the first-order volitions can be considered as “impaired” or “insufficient” ways to show plain consciousness, and second-order volitions can be considered as plain forms of AC. Bringing this together, we might show a way for make sense of what to expect in practical domains. Consciousness and agency might have a natural gap between them.

It is not the case that we should consider the first-order volitions as non-agent related as this would be a tough explanation. Consider hunger and the acknowledgment of being hungry as a motivation for other first-order volitions – could we consider “being hungry” as non-agent related in a hard sense? For Frankfurt, this would clearly bring about a situation in which there is no place for a “responsible” agent. Being hungry is perhaps a conscious state, but not something that characterizes agency in its full sense.

But the same situation, for Block, shows an AC taking place, and certainly PC working. Block could consider an individual as displaying agency when hungry, with this Frankfurtian first-order desire also playing the role of a state poised to an action of looking for nourishment.[1]

The link between AC/PC and agency might not be clear at first glance, but they sure do share some interesting points – for example, intention, in the case of AC and agency, and "having something that is like" in the case of PC and agency. The link might also help us to understand how discussing AC and agency can help untie some misperceptions toward agency as a marker of consciousness.

For example, if we analyze that being poised has to do with the intention of doing something – that it takes someone who is invested in the action – we can understand how AC might stand closer to the concept of agency. It might be true that it takes an agent to have AC**,** but not the other way around. Notice that having AC, being poised in the sense that Block describes, is not per se sufficient for proving someone to be an agent as a person might have the intention to protect themselves against harm (like the case of the blind patient) and still have their agent status questioned. This might also be true in the case of PC, in the sense that PC is per se not sufficient to prove any sort of agency – although they both might share the sense of "what it is like to be something".

Both Frankfurt and Block might agree that having a motivation toward something, and "meaning" it, does not necessarily show agency. For Frankfurt, only persons – those who can have second-order volitions – can have this skill, and for Block it is possible that there would be signs of AC without PC. Akinetic mutism might be a good example to support this, as we will see.

The suggestion that agency presupposes only second-order volitions or "full" AC is not too controversial. This is largely accepted in the realm of philosophy of mind. On the other hand, suggesting that PC or even first-order volitions are something and

[1] One could also claim that the hungry person looking for nourishment displays a rational behavior supported by some rational inference "in his soul", and an inference of a kind that Elizabeth Anscombe envisioned when commenting on Aristotle's view on intentional action (1958): "I'm hungry, and here are some foodstuffs", and after that one sees me eating the food. For Anscombe, here we have the premises (maybe the PC states constituted by my hunger and visions of certain foodstuffs) and a conclusion – that is, an action (that can be publicly observed).

somehow reveal agency is very controversial. Nevertheless, if it is true that AC might be impaired enough to hide itself, as we have seen in locked-in syndrome, then agency is also hard to prove, and maybe not a good way to prove consciousness.

Nevertheless, it seems to be true that consciousness versus agency might work in this direction, but not the other way round:

> The objection is that intentional agency cannot qualify as a marker of consciousness because intentional agency presupposes consciousness. More fully, the worry is that since an agent acts intentionally only if they are aware of what they are doing, it follows that we cannot ascribe intentional agency to an agent without first determining whether or not they are aware of what they are doing. But if we need to do that – the objection continues – then we cannot employ intentional agency as a marker of consciousness; rather, we must instead employ consciousness as a marker of intentional agency. (Bayne, 2012)

I assume the view of intentional agency as a marker of consciousness as doing more trouble than actually helping us to solve anything. Since consciousness is a private experience, how can we be sure of intentionality in cases of disorders of consciousness? In this case, Bayne (2012) is right to assume the opposite – that consciousness can be a marker of intentional agency while intentional agency does not seem to be a good marker of consciousness. Of course, we still have the issue of determining if someone is conscious or not, but in the case that we do find this answer I see no problem in assuming consciousness to be a marker of intentional agency. On the other hand, intentional agency as a marker of consciousness is a different matter – assuming we do not (in the practical field) have the answer of what it is like to be in a disorder of consciousness, it seems quite a big jump to assume intentionality – in this case, this theory of intentional agency as a marker of consciousness only buries the answer we are looking for even deeper.

One suggestive study about the vegetative state and “command following” was published by Colin Klein (2017), and might contain some answers in the discussion of agency as a marker of consciousness. The basic idea was to put a variety of patients diagnosed as being in a vegetative state and search for responses to commands that would (in theory) require intentional agency.

But, if we take Frankfurt's view along with Block's, we can say there is an extra way to go with that, as maybe the first-order volitions can be considered as "impaired" or "insufficient" ways for showing plain consciousness, and the second-order volitions can be considered as plain forms of AC. Bringing this together, we might shine a light on a way of making sense of what to expect in practical domains. There might be a natural gap between consciousness and agency.

> It is not the case that we should consider the first order volitions as non-agent related – but this would be a tough explanation. Consider being hungry, and the acknowledgment of being hungry as a motivation to other first-order volitions: could we consider that "being hungry" is non-agent related in a hard sense? For Frankfurt, this would clearly state a situation in which there is no place for a "responsible" agent. Being hungry is perhaps a conscious state, but not something that characterizes agency in its full sense. But the same situation, for Block, shows an AC taking place, and, for sure, PC working. Block could consider an individual as displaying agency when hungry, with this Frankfurtian first-order desire also playing the role of a state poised to an action of looking for nourishment.[2] The controversy has mainly focused on whether such patients are conscious. This is an important question, but not the only interesting one. Suppose for the sake of argument that these patients are conscious. A mystery remains: how are they conscious? What is it like to be them? An answer, if available, is both scientifically and ethically important. (Klein, 2017)

Klein says that the study's evidence that some patients show signs of positive responses when it comes to command following does not automatically place them in a spectrum of intentional agents. This goes exactly on the path that we would expect as observants of a vegetative-state person – how can we give them an agent status?

Think of Nelson's case again (page 28). We know that Nelson reacts to pain, but is this enough to give Nelson a conscious-like status? Even when he shows signs of decerebration? Most of us would agree that this seems at least controversial. Even if there are signs that Nelson responds to pain – like his heartrate increasing, for example

[2] One could also claim that the hungry person when looking for nourishment displays a rational behavior supported by some rational inference "in their soul", and an inference of a kind that Elizabeth Anscombe envisioned when commenting on Aristotle's view of intentional action (1958): "I'm hungry, and here are some foodstuffs", and after that one sees me eating the food. For Anscombe, here we have the premises (maybe the PC states constituted by my hunger and visions of certain foodstuffs) and a conclusion – that is, an action (that can be publicly observed).

– it is very hard to argue that he has “something that is like” being in pain. Nelson is in a severe vegetative state with signs of decerebration – there is probably nothing that is like being Nelson.

Is there something that is like being in the vegetative state, so far hidden from us in such a way that we cannot imagine what is like? The strong evidence in such studies that points to a kind of “preserved intentional agency” might, in fact, be part of an autonomous and automatic response of parts of the brain that are still active – and that does not necessarily mean that there is actual intentional agency taking place there. Like Nelson, those patients in a vegetative state might react to pain as an automatic response to a stimulus in the brain, like a reflex. This does not mean that they have PC when it comes to pain – if that was the case we would certainly think of this disorder as being not a vegetative one but perhaps a minimally conscious one. Again, this does not seem to be the case.

Because we struggle to understand disorders of consciousness such as Nelson’s vegetative state, and because there are some potential problems with the borders of descriptions, it is unlikely – at least in the present time of this work – that we will pin down what amounts to enough evidence to prove that a disorder *is* or *is not* intentional. Vegetative states, at least for observers, seem to be mental states in which bodily responses are automatic and unintentional. Nevertheless, the label “vegetative” is misleading in Klein’s view:

> I argue that labels like “Vegetative State,” while diagnostically useful, do not pick out natural kinds with respect to consciousness or conscious content. Because of this, we ought to instead focus on what I will call aspects of conscious states. *Aspects are gradable dimensions along which conscious states can vary.* There are potentially many such dimensions. Variation along distinct dimensions can lead to diagnostically indistinguishable states, and variation along the same dimension can put a patient into deferent diagnostic categories. (Klein, 2017; my italics)

People can respond to certain stimuli or orders, and yet this does not classify the actions as “behavioral intended responses” (Boly, 2007). As far as the vegetative state goes, some patients have some responses, such as vocalization responses to pain, while

other do not. Patients who are in a decortication or cerebration level of the vegetative state, for example, might not display any vocal response to pain. We can grade those "patients" in levels – some states being less severe than others, and even temporary.

Imagine, for example, that you are cooking and, not being careful enough, you reach out your hand and get burned, and pull your hand away from the fire. Now, this is tricky because *you might think you wanted* to do that, but, as we know, this is not true. Pulling your hand away is an automatic response that occurs due to a specific defense mechanism named the "reflex arc" (Berkowitz, 2017). This makes you pull your hand away fast enough to prevent damage, and then informs your brain about what happened. There is no decision making or intentional agency here.

It is weird to say that people might have intentionality upon states that are actually "automatic" responses (Klein, 2017), and that this seems to be the case in vegetative states. Hence, the fact that people in vegetative states sometimes responds to some stimulus is not evidence that they behave as agents. For this reason, Klein prefers to use the word "intention" in a loose and broad sense:

> I will use the term [intention] as it is typically used in the literature on impaired states of consciousness: that is to say, rather loosely. In most discussions of impaired consciousness, "intention" simply stands for whatever internal motivational state gives rise to a particular action, subject only to the restriction that it is a succulently complex state that cannot occur completely automatically. Intentional action, in this sense, might (for example) stem simply from some occurrent conscious desire. (…) As such, their responses ought not to be considered as intentional actions. The mental imagery that occurs is, as Zoe Drayson suggests, closer to a "mere happening" than to a full-fledged intentional response (Drayson [2014] pp. 28–9). As such, responses do not indicate consciousness, at least if the route to such indication requires intentional agency. (Klein, 2017)

Nevertheless, even if we use "intentional" with this loose meaning, it does not rule out that there can be something that is like being in a vegetative state (Boly, 2007) (or any other disorder that fails the requirements for full intentional agency). What this state is like or what type of PC they have left is a difficult question that I am afraid there are still no good answers to.

Will Davies and Neil Levy (2017) argue in a opposite way to Klein on this subject, stating that vegetative-state persons have *some* “kind” of consciousness, and we are hence not wrong when we attribute intentionality to them. They argue that - those based studies are persuasive in showing evidence that persons in a vegetative state display communicative behavior. They even say that there is in fact, some “type” of answer being showed in the results – the question lies in what type of answer this is. For Davies and Levy, those patients are somehow conscious. As counterintuitive as this might seem to most of us (me, included), they insist that this is the case.

To construct their argument, Davies and Levy (2017) make use of the following assumptions:

> *Command Following: Command following is a marker of intentional agency. Hence, evidence for command following also provides strong evidence for the presence of intentional agency.*
>
> *Agency: Intentional agency is a marker of consciousness. Hence, evidence for intentional agency provides strong evidence for the presence of consciousness.*

The first assumption is that command following *necessarily* implies intentional agency. This is a conceptual assumption. But, in doing that, the authors *beg the question*, and the argument seems to have a lack of neutrality when it comes to proving its own point. There is no doubt that command following and agency can appear together – this is in fact true in some cases, but in others the bond between command following and intentional agency does not hold or is not strong. If we assume that command following necessarily implies intentional agency we take a conceptual assumption that can conceal an empirical issue. The issue of determining if one has some kind of disorder of consciousness is a clinical issue. The connection between command and intention can be conceptual. Nevertheless, conceptual constraints have consequences for our descriptions. Whatever our conceptual choices, it is important that they are empirically

adequate. That is, they should be apt for making our empirical descriptions finer and more appropriate.

First because spotting command following is not always so clear, and second because we do not have access to mental content, we might be misled in thinking that there is intentional agency when in fact there isn't (Boly, 2007. Klein, 2017). Command following lacks one extraordinarily strong attribute of what we mainly sustain when defending agency – voluntary behavior.

Because command following implies that the subject *is being told what to do*, it is also true that we might be discussing the matter of automaticity and not voluntary agency behavior (Klein, 2017). At this point, another good case to discuss would be that of akinetic mutism. Imagine this:

> ***David's akinetic mutism****: David lies still on the bed. He does not move voluntarily and he does not talk much. Many of his responses – when there are any – are simple and monosyllabic. Nevertheless, he can still talk, but does not do so. He feels hungry and can eat alone, but if not fed will not ask for food. Frequently repeated commands are followed by David, but in a slow, incomplete way, as he lacks something in doing this. David does not have any muscular problems and "shows" sign of being a person of "normal" intelligence.*

Akinetic mutism is in fact a rare disease that has several grades. There are less severe forms – in which the patient is somewhat like David, being responsive – and severe cases in which the patient is not responsive at all, remaining silent and still – such cases often show signs of a bad prognosis. To figure out akinetic mutism, imagine what it is like to have "nothing that is like" being them. This is a hard task, if not impossible. Klein (2017) describes these patients having "a blank in their minds".

Although akinetic mutism might have more than a couple of explanations, this is not important for us here. The important thing here is to understand David's experiences

– if there is something that is like being David and, if not, what is there to learn about in terms of what agency represents to David.

We seem to be able to perceive David's conscious in a way. David has input-output – he responds to questions – but, on the other hand, we cannot claim that David has intention and this is the challenge in claiming David as being an agent. If David has nothing that is like being him, or, in other words, if he lacks PC, we could also ask ourselves if AC without PC exists whatsoever. David seems to lack any sort of qualitative character.

The fact is that, in David's case, there seems to have been an emptying out of content in such a way that he fails to originate any PC:

> One might think that behavior is necessary for consciousness, or that consciousness depends on intentional action and so is abolished if intentions are completely absent. Watt and Pincus suggest that the "emptying out" of consciousness characteristic of AM may lead to the lack of phenomenal content in severe cases, because stimuli must "have at least some potential affective significance in order to gain access to the conscious workspace". (Klein, 2017)

The argument starts to point out to a much weaker relationship between consciousness and agency than expected. The bond between these two depends on how intentionality takes place, not mere stimuli. It is true that consciousness can be present without some strong sense of agency, but nevertheless agency does not seem to be present in patients who lack some sense of consciousness.

The tricky part of all of this is how consciousness can be disguised. Consider this:

> ***Susan's locked-in syndrome case****: Susan is a 30-year-old woman who has suffered a severe brain injury and woke up in a hospital bed without being able to move or speak. She feels everything – pain, emotions, she has thoughts and can reason about things. The doctors assumed she was in a coma, so most of the time people don't refer to her as a "full person" or even talk to her before procedures.*

Susan hears everything and hopes she will get better soon, but she has no way to communicate that she can understand what is going on.

Of all disorders of consciousness, locked-in syndrome holds a very peculiar and interesting status for philosophy. This is a disorder in which a conscious, mute patient is completely paralyzed apart from some form of eye movement, usually as a result of an infarct in the ventral pons. Such patients are often assumed to be in coma and, as a result, may be distressed by inappropriate conversation around the bedside. It is a shell with a pearl inside – but how to make sense of the pearl if we cannot see it?

First, let us begin by stating here that, as for PC, locked-in syndrome must possess this. If we can imagine what it is like to be in a locked-in syndrome, we must agree that there is something that is like being in a locked-in syndrome. There are perceptions, emotions and feelings attached to the being, they are just locked in there.

There is something that is like being Susan.

We think of locked-in syndrome as being with full cognition and mental capacities with no way to interact with the world, with everything else right where it is supposed to be. It is very easy to argue in favor of consciousness with full AC working, but how to prove consciousness in beings who do not “poise” their speech or muscular actions with voluntary behavior?

To begin with, we must agree that having consciousness has nothing to do with muscular voluntary response, which may be a very effective way to prove consciousness but has nothing to do with what we think of consciousness being. People may lack the ability to respond with muscular activity and still have full consciousness – think of a person in a wheelchair, paralyzed from the neck down, for example.

First, we need to understand what the difference in practical terms is of phenomenal without AC, access without PC and disorders that present issues with both.

One might argue that a “zombie-like state” – being conscious without PC – is totally plausible, and compatible with life. A “zombie”, some say, can still have experiences – maybe not the full experience, but some might say that just because the experience is not “full” then a “zombie” would not have any sort of experiences.

Consider a blind patient again – this is a state in which the patient has no phenomenal states upon seeing something, but can still somehow act upon whatever they have has seen. A sort of “zombie”. This patient *might have* a sort of experience upon acting in accordance with what they have “seen” (although they will always claim that they have not seen anything that could lead them to act upon it). Nevertheless, the experience can rest upon many things – their feelings based on their action, their perspective of acting, etc., but **never on the sight of it.** They have **no experiences** of seeing what has caused them to act the way they did. Again, my point that the qualitative perspective of the experience is affected seems to hold strong.

If there is only access, like in the case of akinetic mutism, the qualitative aspects of the experience are certainly affected.

Now, if both phenomenal and AC are affected then we enter a realm where the discussions of value are much more complicated. No one would say it is an ethical issue to allocate resources to a blind patient or one with akinetic mutism, but allocating resources to patients with neither phenomenal nor AC can be rather problematic.

If there are different types of states in which both phenomenal and AC can be affected on different levels, it seems fair to weight them differently too. On the other hand, talking about value in medical ethics considering only one side (the decision-making individuals – medical staff or family members) is very paternalistic and, again, problematic.

As we said above, Block argues for AC as being: “(1) inferentially promiscuous, that is, poised for use as a premise in reasoning, (2) poised for rational control of action, and (3) poised for rational control of speech.” Assuming that Block is right, then (1) is definitely present in locked-in syndrome. The strange part of all this is that we often think

that either a person has “full” AC, or not, but if we think seriously of locked-in syndrome we cannot deny (1) as being present.

If it is the case that a locked-in syndrome patient can have (1) and, let’s assume for a moment, not have (3), there must be different faces or ways for AC to present itself, different abilities and capacities that can coexist or not. We must assume that there can be instances of one without the other.

A locked-in syndrome person such as Susan seems to be an agent – and therefore conscious. In the sense that this person can use their rationality and intentions for decision making and motivate desires and wills, the lack of attachment between mind and body in this case causes a type of disorder that has been constantly proved to be one of the hardest disorders to be in.

Agency and consciousness therefore do not seem to sustain a necessary connection, and agency in several circumstances seems not to be a good marker of consciousness in clinical and practical fields. If there is any good marker of consciousness, agency alone does not seem to a good example. We will have to search for a better answer elsewhere.

CHAPTER THREE

Consciousness as experience

Consciousness goes right through various fields of knowledge and has many definitions. It is a skill not only related to our biological “normal” state, but also where individual and social aspects of consciousness are very important to us. Consciousness is perhaps the most complex human skill that has both a biological and social impact on the way that we live. The fact that we need to understand the science behind the brain is undeniable, but establishing some lines of work even though we do not yet understand consciousness is still much needed.

In chapter two I talked about the role of agency as a marker for consciousness, and the conclusion was that agency must be discarded as a reliable and accurate marker for consciousness. The problem, I believe, lies before the issue of “markers” itself in the methodology we use in the search for markers. Because consciousness is so complex and varied, and because we cannot rely on a simple aspect of it in order to prove its presence, a conclusion is that a marker (if there is one) of consciousness is as complex as consciousness is in itself.

One reason is that, because we do not yet understand how consciousness works, it is very hard to claim a single marker of it – if we do not know what the object we are dealing with is, how can we be sure about a marker of it? Second, because we do not understand agency all that well either, it seems that we are just pairing two “skills” that seem to relate to each other but are not necessarily markers of one another. We constantly struggle to simplify the problem of consciousness when, in fact, there is nothing simple about it!

Agency, as we have seen, can in some cases be a fair marker of consciousness, yet it does not apply to all cases. Why is that? One possible reason is that agency is only one aspect of consciousness besides many others.

There is an explanatory gap when it comes to consciousness (Chalmers, 2010), and to do work that is good enough we need to try and imagine what is lacking from the explanation – even if we cannot get the answer in the present time. Even though we have a practical limit, once we begin to accept that there are no easy answers we can begin to unfold the multiple aspects involved in the topic of searching for consciousness.

First, even if we are still not in agreement that we have a stablished science of consciousness, we have at least some commonsense views on this phenomenon. People, even those who are not into science, perceive and are to some extent capable of recognizing consciousness and "working" with it – making assumptions about its existence in the world, and drawing a line where they think it is over. Think of an uneducated person baking a cake. They know that the cake needs yeast to grow and, through common knowledge and "trial and error", that they cannot open the oven while the cake is baking. They do not understand why and how this process works, but still benefit from it and are able to use it to provide information in decision making and what they think of the world. People have a commonsense view of what is meant by "making a cake" and have an operational approach to it. Something similar happens with consciousness. We do not have a scientific description of what consciousness is, but we have some commonsense views and are apt to use them.

Second, there is a science behind consciousness with medical and measurable aspects, even though this still needs to be developed. As far as science goes, it is known that we do not need full scientific models to affirm that we do know something about one subject – Einstein's theory of relativity, for example, although not yet fully understood, does have some conclusions that we are more than satisfied with when approaching the subject. The need for a full scientific model can be accountable, but this does not eliminate what we do already know about one subject – even if there are pieces still missing from the explanation.

Third, there is a part of the experience of being conscious that will always be only accessible by a first-person point of view. This subjective aspect of consciousness can only be perceived by the person themselves and can only be imagined by the third-person

point of view. “What it is like to be”, in other words, is what we imagine the first-person point of view is, and what mental content it might have.

It is ***not*** fair to say that as scientists we know nothing about consciousness – we experience consciousness, in our own first-person point of view, and all of us have experiences that are derived from consciousness. Having consciousness of our own does not make us an expert in it (Dennett, 1991), but it can give us some clues on what to expect or look for. This is why we investigated some aspects and disorders of consciousness – to have a better understanding of what we know and what is yet missing from the picture. This type of study might help us light the path we still need to walk and also help us discard theories that do not serve us in the practical fields. As we have previously discussed, consciousness and disorders of consciousness might have “something that is like being” that person which might give us an extra clue on the subject if we take the next step. Therefore, if consciousness has something that is like being conscious, then we seem to **experience** consciousness.

One thing we seem to know is that consciousness produces experiences (Chalmers, 2010). Experience is something that we can try to understand and categorize, and we do have fields of science that try to understand experience itself (psychology, for example). Experience seems to be, in my opinion, the core of consciousness and disorders of consciousness – we experience things because we are conscious or have some type of consciousness left. As David Chalmers (2010) would say (agreeing with me, I believe), **the problem of consciousness is the problem of experience.**

While agency (chapter two) does not seem to hold as good an argument for discussing consciousness on every level or every aspect of consciousness (especially for some disorders of consciousness), experience does. On one hand, experience is not a marker (as agency is stated to be) but a *product* of consciousness. On the other, experience is very much attached by “what it is like being in that situation” or “to be that self” – in other words, the conscious *phenomena* that determine what it is like to be an individual.

This has nothing to do with the function itself, and I will try to make an argument for why that is – by looking at the brain and tracing its functions, trying to figure it out to

correlate its behaviors, we can understand and answer questions of function and "what function goes with what behavior." There are some good experiments that help neuroscience understand the problem of how the brain works, but this has to do with functions, not experiences themselves.

Understanding a brain function and its biology or chemistry, or what function goes along with what part of the brain, is a *soft* problem of consciousness. These are, by their nature, purely answerable by experimental studies and completely objective. For example, we already know that emotions are processed mostly in the frontal lobe of the brain – there are no disquieting doubts about that. We also know what damaging that part of the brain may cause, what the clinical signs and symptoms are, and what the diagnosis is. What we do not know is what it is like to be that individual during that particular experience (or why/how they experience things the way they do, if they actually do). As for functions, we know what they are, but consciousness is a different matter.

This has led us to conclude that experience and function must be different matters. With that in mind, some philosophers have made a step beyond, arguing for a new modality of dualism. Chalmers advocates on behalf of an *explanatory gap* between function and experience. This is plausible, for it is extremely hard to qualify experiences based on the function settings of one's brain. Experience, in his words, have no function (2010, p.4). We are empirically able to find correlates for some behaviors in the brain – for example, we know that when we are happy a certain part of the brain is at work, but pointing out correlates does not answer the question of how or why we experience things the way we do.

Although there can be some approximations (and we certainly use them to run diagnoses), the task of binding function and experience is still hard. Some states that seem to share some functions might be "experienced" in a completely different way by different individuals – the answer to this disparity remains, once again, disquieting.

Nevertheless, this represents a problem for the practical fields when it comes to the ethical issues that may arise on account of disorders of consciousness. Since the explanatory gap exists, and since we rely on methods that are not good in judging how "good" the experience is, we fall into a dark zone of ethics.

Compared to the so-called “hard problem of consciousness”, the hardest *ethical* problem of disorders of consciousness is in determining how an individual is experiencing certain aspect of their life and how far this kind of experience is an illness, considering that it is too hard for us to grasp what is it like to be them. The tricky part is that we have no way of knowing for sure right now what is it like to be them, and this presents a problem for asserting that those kinds of experiences are not kinds of wellbeing but of illbeing. Since experience is subjective and since we seem to lack a good way of fulfilling this gap without overly simplifying the problem, we need to discuss how to address this issue in the practical fields. Experiencing seems to be a basic factor when it comes to the discussion of consciousness, and it might lead the way in discussing ethical issues on a fair basis (Chalmers, 2010).

> The really hard problem of consciousness is the problem of experience. When we think and perceive, there is a whir of information processing, but there is also a subjective aspect. As Nagel (1974) has put it, there is something it is like to be a conscious organism. T as the subjective aspect is experience. […] If any problem qualifies as the problem of consciousness, it is this one. In this central sense of “consciousness”, an organism is conscious if there is something it is like to be that organism, and a mental state is conscious if there is something it is like to be in that state. (Chalmers, 2010)

But before answering the ethical problem of how bad one is if one is experiencing a disorder of consciousness, we need to agree about what consciousness is. For since there is an explanatory gap, if we say that one is bad because they are functionally deficient, it is still open on how bad it is for this individual to be in this deficient functional state. Let us try some paths in order to solve this difficulty.

One path is that taken by Chalmers (2010), who suggests that while function is one thing, consciousness must be some different kind of thing – in his words, consciousness must be a *fundamental law* – in other words, consciousness is present in every part of the universe, and all there is possesses some sort of consciousness. A fundamental law is a statement that predicts or best describes a natural phenomenon, in this work consciousness – everything in the universe is ruled by this law, and consciousness is therefore present in every part of the universe and every entity (at

different levels, of course). From a particle to a complex organism, all have some sort of consciousness. A fundamental law should explain why a phenomenon occurs, be successful in repeating its predictions and describe what makes the phenomenon happen (Chalmers, 2010).

This is a smart move, although it is merely conjecture at this point. Think of fundamental laws such as light, gravity or movement. Although we do not know how to address the rules of this fundamental law (consciousness), determining it as such postulates that in order to have experiences an individual **must** have some sort of consciousness.

And, for sure, understanding or knowing that a fundamental law exists does not necessarily mean that we can replicate or answer the question of "how to" entirely. For example, Einstein postulated the rules for the speed of light, and although we know he is right we are not able to replicate this ourselves, and some believe we never will. I agree with Chalmers that this might be the case of consciousness – it may be a fundamental law of the universe that we do not yet know how it works.

This might also suggest that, as a fundamental law, all particles of a living organism might have "some sort of consciousness" – even if it is not a complex one. This is something Chalmers also advocates for. This is a theory called "panpsychism" (Skrbina, 2021. Chalmers, 2010) – pan means "all" and psyche "mind", leading to the view that all things have a mind-like quality. Even particles or very small entities such as atoms have minds – and consciousness – although, of course, not complex ones. Nevertheless, this is not a formal theory, only speculation.

Although this might seem like a very counterintuitive theory because not all things seem to have a mind-like quality (Popper, for example, argued against it in his 1977 work *The Self and Its Brain*), some point out that this type of approach has its pros:

> With the advent of Darwin's theory of evolution in the mid-1800s there came new support for both continuity and non-emergence arguments. If humans evolved from lower animals, they from single-celled creatures, and they in turn from nonliving matter, then the continuity of beings suggests a continuity of the fundamental qualities of experience, awareness, and mind. Evolutionary continuity over time

> makes difficult any attempt to define the supposed point in history at which the mind suddenly appeared. Haeckel (1892) was the first to offer an evolutionary argument, but Paulsen, Royce, Waddington, and Rensch made essentially the same claim. (Skrbina, 2020)

Nevertheless, there seem to be some that see the need to draw a line (Chalmers, 2010) in hierarchical terms – some may have a mind, and some may have only a few aspects of it, leading to the conclusion that not all minds are the same, and some might have a higher level of consciousness (or no consciousness at all perhaps, but just small amounts of it). This does not rule out panpsychism at all (Skrbina, 2021) – it is not an objection – but only states that higher levels of consciousness are dependable on a lower-level element of consciousness. This is not an issue for us in this work since we have already discussed a similar problem (if access consciousness is dependent on phenomenal consciousness and vice versa).

In opposition to that is Daniel Dennett's (1991) path (or one that once took) advocating on the behalf of consciousness as an illusion. He believes that once you explain all the biology, functions, and physical processes you will have explained everything. In his own words, consciousness as we think of it does not exist, thus it is but the result of a physical process – there cannot be hard problems of consciousness.

While studying the nature of consciousness we need to be aware that we are referring to much more than just a biological process that "somehow creates a mind" (although some philosophers like Dennett may dispute this claim) – there are several implications for reducing consciousness to the physical brain's processes, and not solving the problem of consciousness in the practical fields fairly is one of them. How can we be sure that once we explain all the physical processes of the brain we will have explained everything? This seems to leave out many questions that we do take as subjects of both philosophy and the practical fields. Take experiential mental contents, for example – do they have some physical structure? Even if in the future we could describe them as processes with some physical structure, they will continue to not be the experiences as such. As Moore (1902) has argued, this would still be an open question.

Reducing the brain to its physical processes may even rule out believing that consciousness is real (Chalmers, 2010), but how could consciousness be an illusion if we have experiences? How could those experiences be an illusion? In which sense of "illusion"? Surely, as Descartes famously argued, we can have illusions, that is, illusionary experiences, but we cannot ourselves be illusioned that we have them. If it were the case that consciousness is nothing at all then there would not be any questions of the sort of what is it like to be in some state of consciousness, and then disorders of consciousness would not be a problem for ethics – we would not have any issues of how well one is when one is having any experience at all, for they would not be having NOTHING.

Just because we do not have a way to do this now, it does not mean that consciousness is not there at all. Saying that consciousness is an illusion might solve the problem of consciousness for some – if it does not exist then there is no problem to investigate – but this does not seem fair on how many situations one might find themselves in, such as the first-person problem of how bad one is if they experience some disorder of consciousness. It is very noticeable in the practical fields that disorders of consciousness do exist and share a common trait – a reduced level of consciousness. So, if there is something that we see as missing, how can we say it is not there if it does not exist?

Consciousness as a system

Consciousness *might* be the product of a system – like the digestive system or urinary system, for example. The situation of organs working together to produce one thing – consciousness – is again the same as the digestive system, which has lots of organs working to produce one thing – digestion. Understanding consciousness as a body system helps us in many ways –mostly to organize what goes along with consciousness, but also comprehend and point out what to look for when it comes to understanding how consciousness arises – if we know the "map" of it, let's say, it becomes easier to search for the how and why.

The definition of a “system” (Farlex, 2020) in medical terms is:

1. A set or series of interconnected or interdependent parts or entities (objects, organs or organisms) that act together in a common purpose or produce results impossible by the action of one alone.
2. An organized set of principles or ideas.

And:

1. A group of interacting, interrelated or interdependent elements forming a complex whole.
2. An organism or body as whole, especially regarding its vital processes or functions.
3. A group of physiologically or anatomically complementary organs or parts.

This suggestion has many implications:

a) Systems must share some sort of anatomic similarities, even across species.
b) The pathway of the system must have some logic and order, even across species.
c) Similar damages to similar organs or parts of organs must result in similar kinds of illness or disorder.

First, for a being to be conscious (human or non-human) there **must** be some similarities in their anatomic construction. Think of a cat, for instance. A cat has a body in which there is a brain, a spinal cord and so on. Humans also have those same organs, but there are some differences in complexity. From this it seems plausible to infer that the more complex the system, the more complex the experience of consciousness.

What we know about consciousness as a system is that:

a. There is a brain (or at least a rudimentary equivalent thing).

b. There is are input and output biological processes that produce experiences.

It might take a whole organism to produce consciousness, a holistic system, but this is something that is not set as being certain – it is merely just one possibility that might (or might not) add to the answer of this work. Nevertheless, maybe this is why many people fail to see consciousness as a system (or the product of a system). Think of things that may alter our state of consciousness or influence our experiences – hormones, for example. Consciousness might be deeply different for individuals who have vastly different cortisol or adrenaline levels (Shapiro, 2000). This might affect how they experience the world, what it is like to be them, and what type of experience it produces.

This might of course be an issue for our very well-known discussion of the possibility of machines being conscious (Shapiro, 2019), and it might be true that only biological living individuals are capable of consciousness as we think of it, and if there is a machine that can replicate the conscious system *very well* it would still lack something – authentic individualism or a subjective self, in other words.

Lawrence Shapiro (2019) argues a similar thing in his book *Embodied Cognition*, where he discusses themes of embodied mind through many theories, one of which is the dynamical system. One of the views Shapiro presents is that the brain responds to the environment and the internal stimuli sent by different organs and organic components (such as hormones). Machines could not achieve this, and it seems to me a logical conclusion that different individuals and disorders of consciousness (and even the same disorders of consciousness in different individuals!) might be unique.

Having something that is like being that individual from the first-person point of view seems to be only possible if it relies on a unique setting of biology. Machines endowed with artificial intelligence, in this approach, would be conceivable “zombies”, with their working assembled parts exhibiting behaviors similar to conscious beings. How could their secret mental states be revealed? Of course, I am not suggesting that only humans might have consciousness, but I am suggesting that maybe only living organisms can have it.

If the system is not that complex – let us say, a jellyfish – consciousness cannot be complex, or might not even exist at all. I am not disagreeing here with Chalmers (2010), who says that all organisms must have some level or degree of consciousness, but I am suggesting that there might be scales of it when it comes to the complexity of the system.

Because consciousness depends on lots of different variables, and because these variables may be different for each one of us – hormonal levels, for example – it seems fair to say that people can have different kinds of experiences depending on the state of their nervous system, and therefore that we can experience consciousness in different ways depending on whether we are unhealthy or mentally disabled. This is also an answer for why mental contents are different even if the person is a perfect clone (an identical twin, let's say) and is subjected to the same environment. To share the exact same experiences and mental contents, one person would need to be both subjected to the same environment (all the time) and have the exact same organism with the exact same levels of everything – from vitamins to hormones, and so on. This is, of course, impossible. Following this account, if even perfect clones might experience consciousness differently, different species also have types of consciousness that are not the same. Furthermore, it is totally plausible that different disorders of consciousness might experience consciousness in different ways or levels – it does not seem to be a problem to accept this as being true.

This might also be the reason why we have such a difficult time in trying to make exact correlations to functions, disorders and experiences – every person, because of their biology, produce a different kind of outcome – although they might share characteristics, we might never be sure of what it is like to be that person in that particular experience. It is not difficult to understand why there is no good method for evaluating consciousness on a clinical level – if I am right, we can have approximations but never the full answer.

Nevertheless, although the product of the systems might be unique (consciousness), the system must have some similarities in other terms.

If consciousness is a sort of system, then beings that have consciousness must also share some sort of *logical pathway* to follow in order to produce that result – the

result is not the organ itself or something that is attached to a single part of the body, but the result of a coordinated effort to make that system work.

Let us again make a simplified pathway of the digestive system to compare:

(Organism1) Eats – Digest – Defecates.

But when it comes to consciousness:

(Organism1') Input – ***(?)*** – Output (=experience).

The last consideration – that similar damages will cause similar illnesses or disorders – seems to hold, as far as correlation can tell us. Or course, this is an idea of what I think of consciousness being – I may be right, but I may also have only part of the answer. I consider Chalmers as having made a good and fair contribution to this discussion, and his progress also seems to point in a correct way (as of today).

Those accounts give us some ways we can discuss consciousness itself and disorders of consciousness – like Chalmers' ideas on how the problem of consciousness is the problem of experience. Thus they are hypotheses, not definitive answers just yet. These ideas do not solve the problem of the practical field – what this kind of state that people with disorders of consciousness experience is. We are then to investigate what to do with the problem we have, lacking information and needing to make decisions in practical terms.

On the need of a bioethical approach for weighting decisions (values) on consciousness

Although it might be plausible for philosophy of mind not to address the ethical status of many subjects (because some subjects do not really need this discussion), not to address the ethical status of consciousness disorders seems to be not fair to the subject at all. It seems plausible that some of us (philosophers and healthcare teams) should consider ethical aspects of consciousness while studying philosophy of mind and disorders of consciousness – because consciousness has issues that affect practical fields

(healthcare). I really do not see how to detach one from the other (the philosophy of mind and the ethical discussion of the subject). As persons, we intrinsically see value in consciousness (Cox, Steele & Colyvan, 2010) – we give some sort of value to those who possess consciousness. What values these are, how much weight we give them and what they represent for this work are the main questions here, and we will try to address these issues in chapter four.

Consciousness, for example, seems to count heavily in decision-making when it comes to life support being turned off. Consciousness also counts when we consider diagnosing one person with a disorder of consciousness or confer responsibility (moral) to someone. Although it might be easier to advocate on behalf of preserving access consciousness as the only eligible trace for making "the best ethical decision available" for the patient, having only phenomenal consciousness seems to share this assumption in many aspects – access consciousness matters, but phenomenal consciousness also seems to matter a great deal for invoking the best ethical decision available.

What type of value status can we confer to phenomenal consciousness and what weight to give it have many implications for the practical fields, including how much we value those beings (the ones with phenomenal consciousness only) when considering which medical path to follow – remember that this deeply affects the allocation of resources for the patient's treatment, and perhaps the prognosis of the illness.

How to consider those values and what to do? This is not an easy answer – like most of ethical debate answers. Consider a life without phenomenal consciousness first – access mental states that are unaccompanied by phenomenal states. Qualitative questions can be flagged here, mostly because our experiences seem to heavily rely on phenomenal states, and we can assume that we confer value on being able to have phenomenal states over our experiences. We value – and this seems pretty much set – being able to enjoy our experiences, and this seems overly attached to phenomenal states. But how to weight AC without PC? Consider the blind patient here – even if there is nothing that is like choosing an obstacle-free way, how could we say that the outcome of the situation (having chosen the best way) does not have value for them? Of course,

there is an emptiness of PC when it comes to choosing – but the result of the choice itself seems to be valuable for them. This *must* (and does!) have some value.

Persons without full phenomenal states might still have good experiences, and, although the quality of the experiences might seem to suffer a lot from the lack of phenomenal states, it seems very difficult to argue that there is no sort of value there. I am prone to believing that phenomenal consciousness has a value within itself. I am not arguing that a lack of phenomenal consciousness, or parts of it, is an excellent experience or even a fully isolated phenomenal consciousness itself – we have seen, for example, the case of blind patients, and we know that this is a disorder, but even within the disorder we can confer some value to this person's situation.

Nevertheless, qualitative questions over the experience from the first-person point of view of blind patients are hard to answer. They seem not to have something that is like making those decisions based on their own sight. If there is in fact consciousness there, and I believe there is, the experience of it is affected at a level that is very hard to determine from a third-person point of view.

If we take that to the next level and consider someone who has no phenomenal states at all, then there would be no experiences to consider – or some sort of state in which there is nothing that is like having phenomenality toward our own experiences. This depth of emptiness of experiences seems to represent a zero value (or very close to it) to most philosophers who consider consciousness a very valuable skill – for sure, in healthcare terms, this represents a state in which the disorder is so severe that it can affect decision making when it comes to choosing resources to invest in this patient. Should we consider investing in someone who has no experiences or no potential to have experiences in the future? This does not seem plausible if we consider the investment itself, but nonetheless might still hold some value for some. I will try to argue this in chapter four.

Some philosophers argue that the loss of phenomenal consciousness is not sufficient to justify the depreciation of this specific type of consciousness (consciousness without phenomenal consciousness).

> Siewert compares the loss of phenomenal consciousness to the loss of color vision; it is akin to moving from a colored world to a black-and-white world (1998, 323). Perhaps the comparison is an apt one, but it is *very* hard to be sure. We don't lose color vision when we lose phenomenal consciousness. Nor do we lose the ability to experience pleasure and pain, emotions or sounds. We lose the ability to experience these things *phenomenally*, but to what extent that is a significant loss is very hard to judge. (Levy, 2014)

Although this is interesting to discuss, I don't see how this would or could influence the practical ethical fields – disorders of consciousness that lack full or part of phenomenal states, such as vegetative states or blindness, are still illnesses, and we still make decisions that sometimes aren't favorable in the sense of investment. Many countries allow the euthanasia procedure for patients diagnosed with permanent vegetative states, and it is very hard to argue that this decision might be unfair or wrong in the sense that it can bring harm to the patient considering their own point of view.

Maybe this suggests that experiences of *any sort* have a value within themselves, and that we are heavily interested in having experiences that are based on our own phenomenal or access states – I mean, if you were the blind person wouldn't you say that your "experience" had value, even if you do not have phenomenal consciousness of a part of it? Of course, this is still a disorder – we still have a problem with having nothing that is like making decisions based on our own personal view. The emptiness of experiences related with the decision is an issue – but we still seem to give value to the outcome.

Conferring value and choosing a path that is bioethically logical (or at least acceptable) seems to be an rising issue to be solved while discussing disorders of consciousness. Because this would point to a way in which we could solve practical issues with the information we have available, and, I believe, make sense for the patient and their disorders, we can try to make decisions that are both fair and viable. Of course, there is a lot to say in this sense, and now, since we are set to find a way to deal with disorders of consciousness as of today, we will try to investigate the matter of decision making with the information we have about disorders of consciousness.

CHAPTER FOUR

Consciousness and uncertainty: Solving the problem for today

Imagine a world where you are not sure of what to do. Imagine that you have some information, but it is not good enough to reveal the outcome of the situation. The process of decision making, in this case, presents a margin of error that you must find a way to work with. The most likely scenario includes you taking into consideration the ways you can go with this and the chances of being wrong. If the chances of being wrong are too high, you will probably choose a more secure way to proceed (Sandin, 2004; Steele, 2007).

While dealing with disorders of consciousness it becomes clear that we have a real evidence-based problem in medicine-safety parameters – signs and symptoms are often confusing while deciding what is the right diagnosis (Farah, 2013), and, although the diagnosis might become clear with time, the first days and weeks of treatment have a crucial impact on the outcomes that are possible for each patient. Consciousness is a subjective experience and it is therefore exceedingly difficult to pin down the elements needed to understand an experience that is not our own. This is tough in the context of disorders of consciousness. Hence, the diagnosis is usually uncertain, and our considerations of how bad one in such disorders are also out of square. Still, we need to make decisions and allocate resources fairly, and that depends on options that are difficult to choose.

In fact, most clinical practice is based in signs and symptoms and observation of the patient. The decision process often involves (or it should) taking into consideration the course of illness, the prognosis, and, the patient's will – in the case of patients with disorders of consciousness, of course, the patient's will would translate into the course of action that will work as the most suitable and ethical for the patient considering their wellbeing and best interests (to feel no pain, for example), and their family might be invited into this process.

> This informal problem description can be recast, slightly more formally, in terms of three sorts of entities. First, there are *outcomes* – objects of non-instrumental preferences. In the example, we might distinguish three outcomes: either I end up dry and unencumbered; I end up dry and encumbered by an unwieldy umbrella; or I end up wet. Second, there are *states* – things outside the decision-maker's control which influence the outcome of the decision. In the example, there are two states: either it is raining, or it is not. Finally, there are *acts* – objects of the decision-maker's instrumental preferences, and in some sense, things that she can do. In the example, there are two acts: I may either bring the umbrella; or leave it at home. Expected utility theory provides a way of ranking the acts according to how *choiceworthy* they are: the higher the expected utility, the better it is to choose the act. (It is therefore best to choose the act with the highest expected utility – or one of them, in the event that several acts are tied). (Steele, 2020)

Let us imagine this: You wake up and have an important meeting at work. You cannot be late for this meeting, but you also like to save gas money. You have two choices. You can:

(a) Take the state highway

(b) Take a shortcut inside the city

The problems are these:

(a1) The highway takes longer

(a2) The highway spends more gas money

(b1) The shortcut might be jammed

The positives:

(a') The highway is usually free of cars

(b') The shortcut will save you gas money

(b2) The shortcut may be free of cars

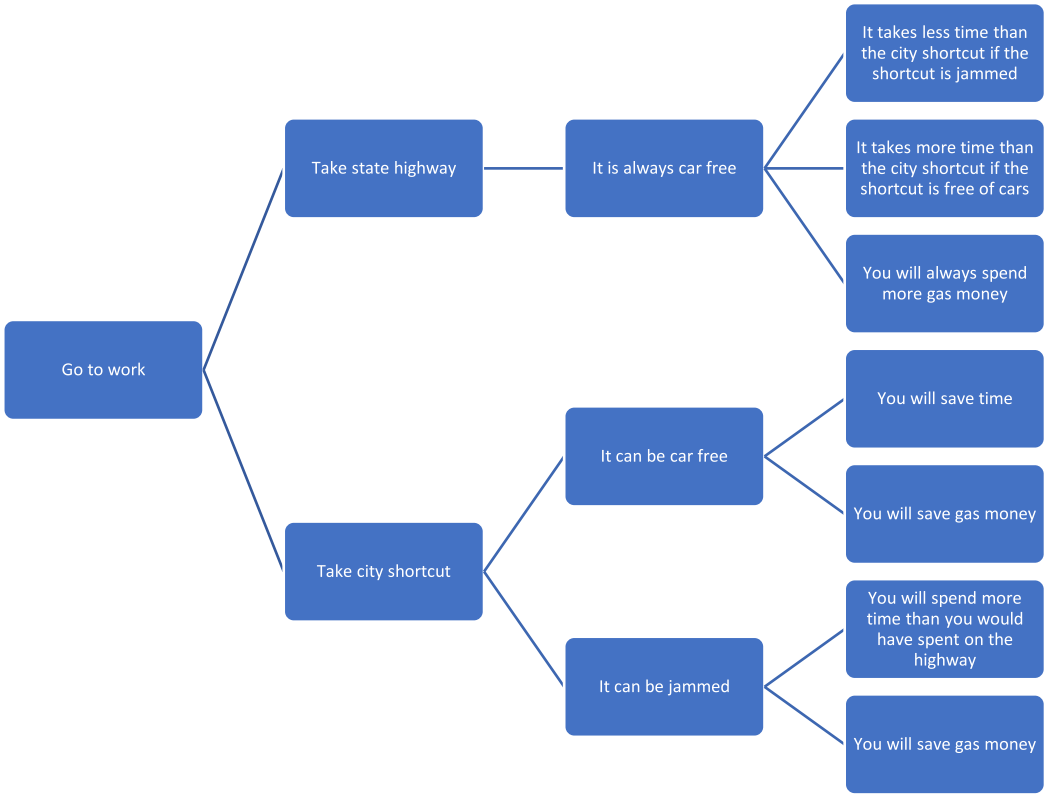
An important observation is that if the shortcut is jammed, the highway will turn out to be faster. So, if you take the state highway:

1) It will take more time than if the shortcut is free of cars

2) You will spend more gas money
3) Nevertheless, the highway is always free of cars

If you take the shortcut:

1) It can be either jammed or free of cars
2) If it is jammed, you will save money, but it will take longer than it would have taken if you took the highway
3) If it is free, you will save money and be also faster than you would have been if you took the highway



Observe that there is a background to this utility calculus – how much you value the items involved in the calculus. If your job matters most to you (suppose you think like this: “If I have no job, I will not have money, therefore I cannot save any money at all”), then your job might have a “heavier” value than money saving. In this case, saving time and not being late to work might have the most important role, which will affect deeply

how you proceed toward the end results of your calculus. In this case, you might turn on the radio and try to gather more information about the traffic to avoid traffic jams in the city shortcut. If you do not have the information you might be willing not to take a chance on the city shortcut (because it might be jammed) and just go for the state highway because you are at least sure that it is free of cars at all times.

Nevertheless, if saving gas matters most for you and you don't mind being late or sitting in your car for a longer period of time (if it will save you money) then you might be inclined to choose the city shortcut regardless. Jeffrey (1983) also has a point in this matter and agrees with this assumption because it goes for the higher expected utility considering the options and their outcomes.

This example might also be applicable in the case of a disorder of consciousness, as in most medical decision making that involves uncertainty. Nevertheless, as seen in the example above, values may vary – in my example I used "not being late" or "saving gas money", but in the case of disorders of consciousness the values might vary from person to person – some might value life itself highly and therefore be prone to consider it worth investing a lot into staying alive no matter what type of experiences are left, and some might be more interested in the experiences themselves and consider that it is not worth investing a lot in a treatment if the experience is not satisfactory.

Considering this individual weighting and how to proceed in the practical fields as professionals, decision-making theories can only play a part here. We cannot postulate a formula that does not consider a high level of variants. For sure, disorders of consciousness do have a lot of variants – what it is like to be in them, what the experience is like, what the prognosis is, what value the patient (or family) considers fair on weighting decisions that will impact the treatment, and so on. Hence, lots of variants, then lots of options and lots of outcomes, and the formula seems to be more complicated than it should be – it does not solve the problem on how to proceed, but it seems to complicate it even more. We are left with another problem to solve (the formula), and we still need to consider a lot of factors. Also, decision-making theory is not, a priori, interested in ethics – which is no problem at all, considering that it is a mathematical calculus, but in the case

of disorders of consciousness it seems problematic. Therefore, I consider that ethics seems to have a part to play here.

Ethics seems to play the role of adding constraints while modelling the decision making. Thus, it seems that the best decision for the patient is not always the best option available – because patients and practitioners might consider different values in weighting options and outcomes. For example, if a patient (or family) considers that not having experiences at their fullest is not worth it – or does not have a heavy value – and decides to not invest in reanimation in the case of their needing one, then does it really matter that we can prolong the patient's life indefinitely? Or that we can have a stable situation with a very low phenomenal consciousness present? Perhaps no. Will the patient be better off if we put them in this situation? If they (and their family) think otherwise, how much welfare can we add to this calculus without being unfair?

So, let us try to use those accounts in our current discussion of disorders of consciousness. Again, the point here is that we must make *decisions under the risk* of being wrong because *we are uncertain of the probabilities.* Nevertheless, if values are, let's say, relative to the agent that makes the choice, as we have seen in the example above, this might also represent a real challenge while applying decision-making theories in the practical field of disorders of consciousness – it is important to mention here that moral issues related to the consciousness affairs have an important part to play, and will establish important boundaries in making this theory acceptable for use.

Making decisions seems to have two important components (Steele, 2010): probabilities and utilities:

> From the formal point of view, these two components play symmetrical roles in decision theory. For each act–state pair (or outcome) we assign a probability and a utility, then we multiply these together and sum the products across each act. The resulting sum we call the expected utility. Standard decision theory then tells us to choose the act with the greatest expected utility (if there is such an act). This is all very familiar. What we draw attention to is that, in terms of the formal decision calculus, probabilities and utilities are treated similarly – they are just two real numbers to be multiplied and then added. (Cox, 2010)

Still, decision-making theories or any theory involving calculations of probabilities (such as Bayesian theories) alone **do not** solve the ethical problem in the practical fields,

and nor does the issue of probabilities. Steele (2010) and Cox (2010) suggest that the question of whether Bayesianism (or even a philosophy of interpretation of probabilities) is really required for the decision theory to work things out is not fair – Bayesianism doesn't add any solutions to this type of discussion. These articles suggest (and I agree) that, despite the symmetry between the probability theory and the utility (Cox, 2010), more should have been written about the ethical side of decision making – this seems very necessary in the practical field.

The rules on how and where to bet are usually proportional on how sure we are of an outcome or how certain we are that we can trust our beliefs – but in disorders of consciousness there is the explanation gap that lowers our probabilities of being right by a lot. I understand that weighting options and outcomes is something that we do all the time – decision making rests on an internal mathematics of reasoning on beliefs – but this alone does not solve the problem of what to do.

If the probability of being right is great you might have a strong argument for acting in a certain way – like my example of the car in the traffic jam where, if you know for sure that the shortcut is free, this seems an easy decision. This seems easy to solve. On the other hand, if you are not sure, the solution does not seem so easy – you might not want to take chances, for example, because you know you can be wrong, or you can gamble and see what happens (Jeffrey, 1983). Gambling on weak probabilities does not seem to be a fair way to approach the issue of disorders of consciousness – there are ethical issues that we need to consider while gambling on healthcare. Is the gambling fair to the patient? Will it produce more good than harm if you are wrong? Again, ethics seems to play an important role here in preventing us from using bare decision-making theories in the practical fields.

First, because when the arguments are weak they can be misleading to the outcomes – and, if we are not sure of the outcome, the range of possibilities can be just too great. Second, if the number of outcomes is too high this seems to drop the "certainty" levels even more.

The best way to solve this, I believe, is to play safe (if possible). If there is too much missing to help us choose a way, we should at least try to minimize damages. Will this

cost a lot in the practical fields, monetarily speaking or in terms of being time consuming? Sometimes, yes. Thus, this seems to be the only way to address the problem without assuming a version of the solution that can potentially cause more harm than good.

On weighing up how to proceed, the gains versus the outcome needs to match our expectations – or at least come as close as possible to it, and, if it is the case, if our level of certainty is low, then the *precaution rule* (Steele, 2006) needs to be applied. Once you have taken measures that rest on the weak argument there is a need to rely on precautions being made so the outcome is not "incorrect" based on the conditions you did not know before.

In doing this, the process of decision making tries to get the best from getting the maximum excepted utility from the situation.

The precaution rule tells us that if we do not have conclusive evidence available then the principle should be used to prevent the possibility of doing more harm than good. This would translate into being cautious in a context of uncertainty.

Although this might seem logical, if we consider that values might be different to different people, it becomes trickier – no matter how great the possible outcomes might sound to one person, if they add nothing to the patient themselves you don't seem to be allowed to perform it.

Placing decision theory and ethics together

Philosophy of mind can only comprehend a share of what consciousness is – this fair share has a direct link to many interesting discussions regarding the recognition of what consciousness truly is. Nevertheless, philosophy of mind and decision-making theories can only go so far. The fact that philosophy of mind deals with one aspect of a thing that is "out there in the real world" makes it, as far as I can tell, something that shares some links with the ethical practical field (or it should). We will try to accommodate decision-making and ethics together in this final section.

First, we need to recognize that philosophy of mind, neurology and neuroethics share a field of work – which, one of many, is the study of consciousness and disorders of consciousness. When philosophy of mind finishes its work, it then begins to find the rest of the answer we are looking for, which is the field of ethicists. Disorders of consciousness have an impact on the practical fields, and practical fields of medicine require ethics (Steele, 2020).

Although very little on the ethical side of decision-making theory has been written, when we are dealing with very subjective aspects of medical problems it seems that ethics has an important role to play in the sense of putting constraints on “rational belief”. Although Bayesianism and other probability and decision-making theories have a lot to say (and sure add a lot to the discussion), they do not even come near to solving the problem of what to do in practical terms (Steele, 2006).

Bradley (2014) argues that ethical uncertainty is often ignored in decision-making theories because they only consider the view that an outcome is valuable if it satisfies maximum utility. Nonetheless, if decision making were to solve the problem of decision making once and for all by doing this, it would not often be misleading when we consider options that are not determined or relative in some sense – and this seems to be the case of disorders of consciousness. We will come back to this soon enough.

One question is how to place ethics with decision-making theories in a way that works in the practical fields, and, despite what I am arguing here, this is already being done, for better or worse. Although it seems obvious that some type of ethical work in philosophy needs to be done in tracing decision-making theory to practical medicine (there are some examples of work in this area but they are very limited, such as the work of Cox [2010] and Steele [2010] that I often quote here), there is much to be said about how to do this. Steele (2010) argues for the need of a type of decision making that runs on *two types of utility*, one of them being ethical value. Ethics here will *demand* a moral constraint in the decision-making process.

> An outcome, Oij, is the result of the agent performing action ai while the world is in state sj. (We will use this notation for acts, states, and outcomes throughout the

> rest of the paper.) There will, in general, be a range of possible outcomes of performing ai. Two constraints on a utilitarian utility function are:
>
> (U1) If Oij involves greater total welfare than Okl, then any admissible utility function u must be such that u(Oij) > u(Okl).
>
> (U2) If Oij involves the same total welfare as Okl, then any admissible utility function u must be such that u(Oij) = u(Okl). (Cox, 2010)

Cox (2010) argues that the constraints presented by him are enough to prevent harm, and that they are (very much) needed to do a fair job in ruling out unreasonable unethical functions that value things that are not logical, like genocide or pain. These can never be part of an ethical possibility.

Of course, despite these clear unethical issues, the type of constraint we are talking about is deeply subjective, along with the type of moral theory we might add to this – which makes the job of dealing with the matter of disorders of consciousness difficult to solve in the practical fields, but also fairer. It is also important to note that my point here is not to adopt some theory of ethics that I think is suitable for solving the problem or defend it – moral duties or prohibitions do not need to be absolute, but rather conditional (Steele, 2007) to each case. Disorders of consciousness vary widely according to type, signs and symptoms, and of course and outcome, so the decisions should only vary as much as this. Therefore, it does not matter to my conclusion if one decides to use utilitarianism or any other account as this does not affect the central point of this work.

So, how to make decisions when we are not certain of someone's condition? Some philosophers base this on the maximum expected utility rate. Expected utility is, according to Jeffrey (1983), using the option that increases the utility of something – in other words, this option needs to equal or surpass the utility of every other option available. Each option needs to be weighted, then the utility is a matter of the weighting of the option over the utility the outcome will have, and this is needed for each and every option since we are not sure about (remember we are talking about disorders of consciousness), so we make a sketch of every outcome possible and weigh our possible courses of action in each, even considering uncertainty – the one with the higher utility should be the one chosen (Jeffrey, 1983).

Although this might seem to be only an example, this type of situation is quite common in healthcare – we usually rely on the odds (probability) of something to make decisions. Nevertheless, there is the problem of weighting uncertainty, which seems to be a great part of the disorders of consciousness problem.

Richard Bradley has also argued in a similar way (2014), but I believe he was more effective in the applicational usability of his argument when we consider disorders of consciousness. He claims that, although decision-making theories help us formulate the problem as in my example above (or in any other matrix-like scheme), this alone does not solve the problem. First, because framing the problem in this particular way only tells us what the maximum expected utility is – which is still dependent on value weighting, and it alone does not guarantee that the agent will always do so. Just because the matrix points in one way, it does not seem plausible to say that this option will be the one chosen or that it is as "available" as the other options displayed. For example, a matrix might suggest that one medication is a better choice over other medications, but can cost four times more or not be available for purchasing at the time the decision needs to be made.

Thus, Bradley also considers that the problem itself might not be representable in any kind of matrix (at least not fairly) because there can be multiple elements that are not traceable while assembling the matrix in the first place (Bradley, 2014), and this might also be true for disorders of consciousness. The consequences (outcomes) might be multi-dimensional, according to the author – the outcome might cross different times and also have multiple future benefits and costs, none of which are very clear at the time when we are assembling the matrix:

> When we assume that values are given, we take this uncertainty to have been resolved in some way. This could be because we assume that there is a fact of the matter as to how good a consequence is or as to whether one outcome is better than another. But it could also be because the description of the decision problem itself comes with values "built-in" (…) In many situations, however, values are not given in any of these ways and the agent may be uncertain as to the value she should attach to the relevant prospects. (Bradley, 2017)

In these circumstances the utility that the agent assigns to a consequence will reflect a subjective value judgement expressing their ethical uncertainty.

One other strong argument presented by Bradley is that the world (especially in medical terms) is, despite our best evidence, not purely deterministic, leading us to the problem of not having a fair predictable consequence to take into consideration. Even if the outcome is set there are too many subjective values that need to be taken into consideration in the medical field – Bradley refers to this issue as a "dependency hypothesis", suggesting that even determined outcomes are not as radical in the need for choice as we might have previously thought they ought to be.

This seems more reasonable in terms of application – radical or purely mathematical representations do not add to the solution if the variants are too numerous. In the case of disorders of consciousness – because we can, in certain cases, lack too much information – I believe that the matrix is not even possible, as Bradley (2014) suggests. In this case, the multi-dimensional scenario is very clear – too much to take into consideration could suggest we could be better off not building the matrix itself but rather taking desirability into consideration.

Desirability is questioned by pure mathematical schemes – many decision-making theories defend that it would be incoherent to choose an option that does not satisfy the best expected maximum utility, but this seems too strong. I agree with Bradley (2014, 2017) that this is rather arbitrary. The reason I say this is because of the conditional aspect of the normative sentence being described by the matrix's best option – which may or not satisfy the chooser in terms of values. In this sense, any choice would be allowable and also logical. If the person building the matrix puts too much value into one option it will automatically weight it more than it would have otherwise, making the choice itself a logical one:

> There are various ways of giving substance to the notion of being preference-based. On an explanatory reading, it means that the decision maker's preferences explain the choices that she makes by providing the reasons for them. On the other hand, on a normative reading, it means that the decision maker's preferences rationalize or justify the choices that she makes. (Bradley, 2014)

The conclusion is that preference (or desirability) is a non-determined issue. In other words, it is relative to one individual's value, given the set of situations presented. What is chosen cannot pass as incoherent if it satisfies the relation of *utility x values* being

taken into consideration. Agreeing with the Bradley, I would also say we need to be vigilant in not giving too much importance to utility maximization in the sense that it is only a validation driven by an agent's preferences, a mere mathematical representation. The situation will (considering disorders of consciousness) be time and case-related pertinent, and for this reason it is impossible to create an effective general matrix at this time (if a matrix is possible at all).

If values are relative to each case and person, the line of treatment chosen to address each disorder of consciousness should also be relative – even if two people have the exact same disorder, it is totally plausible that we will end up addressing each particular treatment in a totally different way. The normative aspects of this choice are much less settled and dependent on each case scenario.

Therefore, to respect someone, our goal cannot be to find a unique treatment according to a certain state of affairs, but to accept not only that each individual is different and has a particular disorder, but also that they have a set of personal values that delineate what is in their best interest (even if that individual is not conscious in its fullest or most demanding sense), a set of personal values that restrict us in what we can or cannot do with that person. More objectively, the overall welfare of one person is strictly conditional on their particularities.

We need to remember that neuroethics and issues involving discussions of disorders of consciousness are rather permeable, mainly because science, philosophy of mind and ethics need to walk together to obtain this answer:

> And in this case the horse is out of the barn: It is clear that interesting and significant work is being pursued regarding the brain bases of ethical thought and behavior and that this theoretical understanding has influenced, and has the potential to influence, our own thinking about ethics and our ethical practices. That neuroethics exists is undeniable. Neuroethical lines of research have borne interesting fruit over the last 10–15 years. (Roskies, 2016)

Thus, it seems that even if the answer is not yet available when it comes to disorders of consciousness and consciousness itself, it seems plausible that we can address the issue using the information we have available, while also being fairly cautious

about making decisions, hence the importance of constraints of ethics over decision-making theories to determine safety limits upon what can be done. We ask too much of decision-making theories in the case of total ignorance or ambiguity – thus, the evaluation of preference considering the likelihood of the outcome is not determined (Bradley, 2014).

Any sort of interpretation of probabilities is not enough to solve our problem, I conclude. If the answer is there (and I believe it is out there somewhere), it surely must include science, ethics and a good theory of philosophy of mind that makes a fair account of what disorders of consciousness are and what is fair when considering personal values and best interests.

FINAL REMARKS

Although the conclusion of this work will lead me toward saying that there is still much to be done in the field of disorders of consciousness, my main conclusion is that we can do a fair job in the practical fields considering the amount of information we do have available today. The gap between what we have as the best explanation available and the actual answer for consciousness is there still, but there are – and I include myself in this consideration – many working on a solution. There will be an answer in the future, I am very optimistic about this.

I believe, as I considered in chapter one, that there is still confusion in diagnosing disorders of consciousness – which, as I pointed out, might affect the prognosis of the disease, since the prognosis depends deeply on the amount of resources invested in the first days of treatment. Nevertheless, I also believe that I have helped to solve this problem in chapter four by introducing the precautionary principle. I have to be fair here and point out that I know that a large amount of work needs to be done to make my statement usable in the practical fields. Nevertheless, it looks like the only way (at least to me) to treat someone fairly and give them “the best and fairest” treatment is to evaluate case by case – there will be no rule to follow.

In this sense, the confusion in the diagnosis itself is still there, but we can try to minimize the damage done by misleading assumptions over disorders. The goal of minimizing the wrongness that can affect the prognosis seems to be the fairest and only possible way to proceed, if we consider how much is still left to be done into the field of consciousness and disorders of consciousness. I am not saying that this solves all problems – I agree that we still need a better way to diagnose disorders of consciousness – but, in today’s possibilities, since we lack this part of the answer, I believe this work responds fairly to the question of what can be done in the practical fields, even if it only minimizes damages and evaluates case by case.

There is some good work available that tries to describe and discuss consciousness like that by Ned Block, Tim Bayne, David Chalmers and Tyler Burge, all of whom have

had an incredible influence on this work. The discussion of access and phenomenal consciousness, followed by the insights on Bayne's and Burge's work in chapters one and two, led me to the conclusion that, although Block did a good job, the acquaintance on the subject of access and phenomenal consciousness has evolved enough to overcome his work – Block's work was a first step, but Bayne's and Burge's are the following steps.

Although I would love to present a general rule to study or find consciousness itself – if there is any, I still lack this answer – it seems that every rule or marker that I have tried to investigate does more harm than good in the practical fields, and does not seem to solve the problem at all in the realm of philosophy. Agency, as I said in chapter two, although used many times as a fair marker of consciousness, is not fair at all! This is not a good marker, simply because it fails for many disorders of consciousness. If agency and consciousness pair up in some way it is not as markers of one another – just because these two have something to do with one another, it does not mean that one can fairly explain or serve as a marker of another, and this is surely the case of agency.

I also tried to investigate consciousness through the views of Chalmers and others (like Lawrence Shapiro, 2019) in chapter three to try and find common ground to discuss the disorders of consciousness and also present which hypothesis I believe works the best as of today. Experience, for me, seems to be the core of the problem when it comes to disorders of consciousness – because we are so interested, as individuals, in having experiences, we value consciousness a lot. I know that chapter three has many hypotheses, none of which are confirmed, but discussing the subject, even if it leads us to no answer now, still helps us take small steps forward, and seems a good way to get closer to the answer.

In chapter four I tried to come to a conclusion in what to do in the practical fields with what information we have available today. Although decision theories such as Bayesianism are very interesting, they do not help us a lot in the case of disorders of consciousness. Ethics take over decision theory at a certain point, and there is a clear need to put constraints on what can be done using decision theory. Why? Because we are dealing with a practical field that deeply affects people's lives and expectations, and

mathematical schemes working alone can therefore only do so much. More needs to be done on the ethical side of decision theory. This seems to me the only way to measure our expectations against reality, and in this matter any decision theory – Bayesianism included – has limits on what it can do without doing harm in the practical fields.

I assume this will be part of my next step as a researcher of philosophy and a nurse – to write about the ethical side of decision-making theory, and continue to investigate the issues of how those can affect and help is in the practical fields. There is a lot to be done – studies on consciousness, disorders of consciousness and other subjects associated with those discussions will probably go on for decades to come. Nevertheless, I assume that this work makes some important steps in how to deal with disorders of consciousness and consciousness itself – if you cannot come up with a fair answer, precautions are needed.

REFERENCES

ASHBY, D. SMITH, A.F.M. **Evidence-based medicine as Bayesian decision-making. Statistics in medicine.** 19. Pages 3291–3305. 2000.

AZEVEDO, M.A.O. **The precautionary principle: and some implications of its use on the risk and safety of new biotechnologies and human body reengineering.** Biotechnologies. Pages 235–274. 2012.

BAYNE, T. **Agency as a marker of consciousness.** AndyClark_09.indd. 2012.

BAYNE, T. SHEA, N. **The vegetative state and the science of consciousness.** Brit. J. Phil. Sci. 61. Pages 459–484. 2010.

BAYNE, T. HOHWY, J. OWEN. A.M. **Are there levels of consciousness?** Trends in cognitive sciences. CellPress. Vol. 20, No. 6. June, 2016.

BERKOWITZ, AARON L. **Lange Clinical Neurology and Neuroanatomy: A Location-Based Approach**. 1st Edition. Kindle Edition. Lange. 2017.

BLOCK. N. **On a confusion about a function of consciousness.** BEHAVIORAL AND BRAIN SCIENCES. VOL. 18, No. 2, Pages 227–287. 1995.

BLOCK. N. **Consciousness.** In R. Gregory Ed. Oxford Companion to the mind. Second Edition. 2004.

BOLY, M. **When thoughts become action: An fMRI paradigm to study volitional brain activity in non-communicative brain injured patients**, Neuroimage, 36(3), pp. 979–92. 2007.

BURGE, T. **Foundations of mind. Philosophical essays.** Volume 2. Clarendon Press. 2007.

BRADLEY, R. **Decision theory: A formal philosophical introduction**. London School of Economics and political science. 2014.

BRADLEY, R. **Decision theory with a human face.** Cambridge University Press. 2017.

BRATMAN, M.E. **Intention, Plans, and Practical Reason.** Harvard University Press. 1987.

CHALMERS, D. **The character of consciousness**. Oxford University Press. 2010.

COX, D., STEELE, K., COLYVAN, M**. Modeling the moral dimension of decisions**. *Noûs* 44, 503–529. 2010.

DAVIS, W. LEVY, N. **Persistent Vegetative State, Akinetic Mutism, and Consciousness**. In W. Sinnot- Armstrong Ed. **Finding Consciousness: The neuroscience, Ethics, and Law of Severe Brain Damage**. Oxford. 2017.

DENNETT. C. D. **The Intentional Stance.** MIT Press. 1991.

FARAH, M. CHATTERJEE. A. **Neuroethics in practice: Medicine, mind and society**. Oxford University Press. 2013.

FARLEX, **System.** Entry. The free dictionary. 2020. At https://www.thefreedictionary.com.

FELDMAN, F. **Pleasure and the good life. Concerning the Nature, Varieties and Plausibility of Hedonism.** Oxford University Press. 2004.

FRANKFURT, H. **Freedom of will and the concept of a person.** The Journal of Philosophy 68 (1): 5–20. 1971.

HOSPITAL SÍRIO-LIBANÊS. Protocolo gerenciado de acidente vascular cerebral (AVC). Documento Operacional. Protocolo Sirio-Libanes. HSL-PROT-CORP 007/REV.02. 2018.

JEFFREY, R C., [1965], **The logic of decision**, second edition. University of Chicago Press. 1983.

KAVANAUGH, J. **Who count as persons? Human Identity and the ethics of killing.** Georgetown University Press. 2001.

KLEIN, C. **Consciousness, intention, and command following in vegetative the state**. The British Journal for Philosophy of Science. 68. 1st Edition. 2017. At http://colinklein.org/papers/VSPaperFinalForWeb.pdf

LANGE, **Clinical Neurology**. McGraw-Hill Education/Medical. 10th edition. 2014.

LEVY, N. **The value of consciousness.** Journal of Consciousness Studies 21(1–2):127-138. Pubmed. 2014. At https://www.researchgate.net/publication/262022370_The_Value_of_Consciousness

LIBET, B. **Unconscious cerebral initiative and the role of conscious will in voluntary action**. In Neurophysiology of consciousness. Contemporary Neuroscientists. Bikhauser. 1993.

MILLICAN, P. WOOLRIGDE, **Them and Us: Autonomous Agents In Vivo and In Silico.** In Alexandru Baltag and Sonja Smits Eds. Johan van Behthem on Logic and Information Dynamics, Springer. Pp 547–567. 2014.

NAGEL, T. **What is it like to be a bat?** The Philosophical Review LXXXIII, 4 (October 1974): 435–50.

NOZICK, R. **Invariances.** PG. 92-93. Harvard University Press. 2001.

PARFIT. D. **Reasons and persons.** Oxford University Press, Clarendon Press. 1984.

RESNIK, M. D. **Choices: An Introduction to Decision Theory.** NED – New edition ed., University of Minnesota Press, 1987. JSTOR, www.jstor.org/stable/10.5749/j.ctttshgd.

ROSKIES, A. **Neuroethics**, The Stanford Encyclopedia of Philosophy. Edward N. Zalta Ed. 2016. available at https://plato.stanford.edu/archives/spr2016/entries/neuroethics

VAN GULICK, Robert**, "Consciousness".** *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta Ed. https://plato.stanford.edu/archives/spr2018/entries/consciousness.

SANDIN, P. **Better Safe than Sorry: Applying Philosophical Methods to the Debate on Risk and the Precautionary Principle.** Stockholm (2004), available at http://www.infra.kth.se/~sandin/dissintro.pdf.

SCHLOSSER, M. **Agency.** The Stanford Encyclopedia of Philosophy. 2019.

SHAPIRO, L. **Embodied cognition.** Routledge. 2nd edition. 2019.

SHAPIRO, L. **Multiple realizations.** The Journal of Philosophy, Vol. 97, No. 12. pp. 635-654 2000. Available at http://www.jstor.org/stable/2678460.

SKRBINA, D. **Panpsychism**. Internet Encyclopedia of philosophy: A peer-reviewed academic resource. 2021. Available at https://iep.utm.edu/panpsych/#H3.

STELLE, K. **What are the minimal requirements of rational choice? Arguments from the sequential-decision setting.** *Theory and Decision* 68, 463–487. 2010.

STEELE, K. ORRI. H.S. **Decision Theory**, The Stanford Encyclopedia of Philosophy. Winter 2020 Edition, Edward N. Zalta Ed. 2020. Available at https://plato.stanford.edu/archives/win2020/entries/decision-theory.

STEELE, K. REGAN, H.M. COLYVAN, M. BURGMAN, M.A. **Right Decisions or Happy Decision-makers?** *Social Epistemology* 21, 349–368. 2007.

STEELE, K. **The precautionary principle: a new approach to public decision-making?** *Law, Probability and Risk* 5, 19–31. 2006.

STEWART, R**. "Environmental Decision-Making Under Uncertainty".** Res. in Law and Econ., Vol. 71 (2002), p. 76.

WILDAVSKY, A. **"Searching for Safety"**. Transaction Books. 1988.