UNIVERSITÉ PARIS-SORBONNE

ÉCOLE DOCTORALE V

Concepts et Langages

T H È S E

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ PARIS-SORBONNE

DOCTORAT EN PHILOSOPHIE

Présentée par :

UNIVERSITY OF GLASGOW

SCHOOL OF HUMANITIES
COLLEGE OF ARTS

Philosophy Department

T H E S I S

in fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (PhD)

Presented by:

Abraham SAPIEN CORDOBA

January 2018

The Unpleasantness of Pain

Supervisors:

M. Daniel ANDLER – Professeur Émérite, Université Paris-Sorbonne
M. David BAIN – PhD, University of Glasgow
M. Michael BRADY – Professor, University of Glasgow

Members of the jury:

M. Murat AYDEDE – Professor, University of British Columbia
Mme. Jennifer CORNS – PhD, University of Glasgow
Mme. Frédérique DE VIGNEMONT – PhD, Institut Jean-Nicod

To my previous (suffering) self,

ABSTRACT

In this thesis I provide an account of the unpleasantness of pain. In doing this, I shed light on the nature of pain and unpleasantness. I propose to understand the unpleasantness of pain based on the determinable-determinate distinction. Unpleasantness is a determinable phenomenal property of mental states that entails badness. I propose that an unpleasant pain experience has two phenomenal properties: i) the phenomenal property of being a pain, and ii) a phenomenal determinate property (u1, u2, u3, etc.) of the unpleasantness determinable. According to this theory unpleasant pains feel bad, and this explains why we are motivated and justified in avoiding them. This explains, for example, why we are motivated and justified to take painkillers. This theory allows us to account for the heterogeneity of unpleasantness, i.e., we can explain how different unpleasant experiences feel unpleasant even if they feel so different.

The thesis is organised into seven chapters and divided by three main themes: i) what the unpleasantness of pain consists in, ii) how we can account for the great phenomenal diversity among experiences of unpleasantness, and iii) which cases suggest that there could be pains that are not unpleasant. Broadly, the first two chapters deal with the first theme, where I analyse two reductive accounts of unpleasantness: the content theories and the desire theories. I deal with the second theme in the third and fourth chapter, where I analyse different theories that try to account for the phenomenal property of unpleasantness. In the fifth and sixth chapter, I focus on the third theme, where I consider different cases that suggest the existence of pains that are not unpleasant. In the final chapter, I offer a conclusion of the three main themes by providing my own view on the unpleasantness of pain.

TABLE OF CONTENTS

ABSTRACT 3
LIST OF FIGURES 6
ACKNOWLEDGEMENTS 7
DECLARATION 8
INTRODUCTION 9
CHAPTER 1: CONTENT THEORIES……..15
1.0 Introduction…….15
1.1 Representational theories……17
1.1.1 Damage account of pain (DAP) ……19
1.1.2 Bain's account of pain (BAP) ……...25
1.2 Imperative theories……..33
1.2.1 Martínez' account of pain (MAP) ……35
1.2.2 Klein's account of pain (KAP) ……..38
1.3 The shooting-messenger problem……...40
1.4 Solutions to the messenger-shooting problem……44
1.4.1 Pain as instrumentally bad…….45
1.4.2 Pain as bad in virtue of other mental states……...48
1.5 Conclusion……..52
CHAPTER 2: DESIRE THEORIES……..53
2.0 Introduction…….53
2.1 Pain's unpleasantness in terms of desire……54
2.1.1 Sid-desires…….54
2.1.2 The heterogeneity problem……61
2.2 Problems for desire theories……65
2.2.1 Unconscious unpleasant experiences……65
2.2.2 The Euthyphro Dilemma……...69
2.2.3 The evolutionary explanation ……...73
2.3 Conclusion……...…77
CHAPTER 3: THE DISTINCTIVE FEELING THEORY……...…78
3.0 Introduction…….78
3.1 The distinctive feeling theory…….79
3.2 Heterogeneity problem for the distinctive feeling theory……...…86
3.2.1 A possible solution for the distinctive feeling theory?.....91
3.3 Conclusion……..93

CHAPTER 4: THE DIMENSIONAL THEORIES OF UNPLEASANTNESS……...95
4.0 Introduction…….95
4.1 Hedonic dimension theory……...…96
4.1.1 The essentiality of unpleasantness…….97
4.1.2 The heterogeneity problem for the hedonic dimension theory……..98
4.2 The solution to the heterogeneity problem……...102
4.2.1 The dimensions of unpleasantness……..106
4.3 Pains that are not unpleasant……108
4. 4 Conclusion……...109
CHAPTER 5: ARE PAINS ALWAYS UNPLEASANT? ……...…110
5.0 Introduction……...110
5.1 Non-hedonic pains……..……..111
5.1.1 Congenital pain insensitivity……..……112
5.1.2 Acquired pain insensitivity……...…115
5.1.3 Pain indifference due to surgery…….…….117
5.1.4 Pain indifference due to lesion: pain asymbolia……..122
5.2 Conclusion……128
CHAPTER 6: MASOCHISM……129
6.0 Introduction……...129
6.1 Means-end masochism…….…….130
6.2 Side effect pains……134
6.3 End-in-itself masochism……...138
6.3.1 Contextual masochism……138
6.3.2 Masochistic pleasures……..……142
6.4 Conclusion……145
CHAPTER 7: THE HEDONIC DETERMINABLE-DETERMINATE THEORY..146
7.0 Introduction……...…146
7.1 The hedonic determinable-determinate theory (HDDT)……...…147
7.2 Two final problems for HDDT……151
7.2.1 The different ways of being unpleasant……...152
7.2.2 The intrinsic badness of unpleasantness.....157
7.3 Conclusion.....163
CONCLUSION.....164
RESUME SUBSTANTIEL.....168
BIBLIOGRAPHY.....203

LIST OF FIGURES

Fig 1: Magritte painting permeation……..89

Fig 2: Magritte painting filter……..89

Fig 2: Magritte painting tacked on……....89

ACKNOWLEDGEMENTS

I wish to thank my supervisors, Daniel Andler, David Bain, and Michael Brady, for their enormous support during the last years; they have been a great inspiration and have taught me how to do philosophy. I want to thank them for their intellectual guidance throughout this process. I also want to thank Jennifer Corns; the final product of this thesis has greatly benefited from our discussions. I would also like to thank Manolo Martínez for our philosophical exchange, and for welcoming me in Barcelona. I am grateful to Frederique de Vignemont for her continuous support before and during my PhD. It is thanks to her that I was able to come to Glasgow and be part of the Value of Suffering project, from which I have learned so much in a professional and a personal level. Throughout my PhD, I have presented parts of my research in several conferences and seminars. I would particularly like to thank the audiences in the Philosophy Department at the University of Glasgow, the Philo Doctes association in Paris Sorbonne, and the *Mente-Investigación* group that we developed in Mexico City.

I have been funded during my PhD by the John Templeton Foundation, as part of the Value of Suffering project, and by the Mexican National Council of Science and Technology (CONACYT), to be part of Paris Sorbonne. These scholarships have made it possible for me to develop a joint PhD between the UK and France. I wish to thank everyone who has been involved in achieving this joint effort. Thanks to the MacRobertson postgraduate travel scholarship, from the University of Glasgow, and from the *Mission de doctorants à l'étranger* scholarship, from Sorbonne University, I was able to visit the Logos Research group and organize a workshop in Barcelona, Spain, during the 2016 summer.

Finally, I would like to thank Gloria Andrada, Antonio Aguilar, Umut Baysan for their friendship, and my parents, Diana Córdoba and Salvador Sapién, for their unlimited love and support, and for staying with me late until night so I could finis*h* my homework.

DECLARATION

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree in any university.

Signature:

Printed Name: ABRAHAM SAPIEN-CORDOBA

INTRODUCTION

Pain is unpleasant. Given that pain is the paradigmatic example of an unpleasant experience, I aim to shed light on what pain and unpleasantness are by trying to understand what it means for a pain to be unpleasant, what the structure of unpleasantness is, and by tackling several problematic aspects of the relation between pain and unpleasantness. By doing this, I will also provide a general account of what it means for an experience that might not be a pain to be unpleasant. Here are some of the main questions that I address regarding pain and its unpleasantness:

- What does pain consist in?
- What does it mean for pain to be unpleasant?
- How can we account for the diversity of unpleasant experiences?
- How can we explain that the unpleasantness of pain motivates action?
- How can we explain that the unpleasantness of pain justifies action?
- Why do we sometimes seek out pain experiences?
- Is pain always unpleasant?

All of these questions regarding the nature of pain and unpleasantness will be developed and answered in detail. To this end, the thesis is organised into seven chapters and divided by three main themes: i) what the unpleasantness of pain consists in, ii) how we can account for the great phenomenal diversity among experiences of unpleasantness, and iii) which cases suggest that there could be pains that are not unpleasant. Broadly, the first two chapters deal with the first theme, the third and fourth chapter with the second theme, and the fifth and sixth chapter focus on the third theme. In the final chapter I offer a conclusion of the three main themes by providing my own view on the unpleasantness of pain. According to this account, an unpleasant experience is something felt, it is a phenomenal property of mental states, and this property should be understood using the *determinable-determinate* distinction. Here is a general description of the content of each of the chapters.

In the first chapter I explain the *content theories*. This approach appeals to the notion of mental content in order to account for the phenomenology of the unpleasantness of pain. Given the explicative power of this approach, it is not a surprise that content theories have given an account of pain and its unpleasantness. I explain the development of content theories and discuss the two main and more successful accounts that attempt an

explanation of pain's unpleasantness in terms of mental content: i) representationalism, and ii) imperativism. Whereas the first view focuses on *indicative* content, with a special emphasis on evaluative content, the latter focuses on *imperative* content. These theories argue that we can explain pain's unpleasantness by appealing to the type of mental content that constitutes the experience.

The main problem for content theories is the messenger-shooting problem. If we accept that the unpleasantness of pain is bad in itself, content theories are unable to explain this feature by appealing to the mental content that constitutes such unpleasantness. In other words, content theories cannot explain why we have good reasons (i.e., why we are justified, why it is rational, why it is desirable) to perform certain actions when we experience unpleasant pains in virtue of such unpleasantness. Content theories cannot explain why we have good reasons to perform actions directed at the unpleasantness of pain itself. To put it simply, content theories cannot explain why it is rational, for example, to take painkillers. Whereas it is rational to take painkillers - it is desirable to do something to stop feeling an unpleasant pain in virtue of such unpleasantness - content theories are committed to saying that taking painkillers is non-rational, i.e., content theories are committed to saying that the content of our unpleasant pain does not provide a good non-instrumental reason to take a painkiller.

In Chapter Two I focus on *desire theories*. According to these, the unpleasantness of pain consists in having an intrinsic desire directed at a pain sensation, where this sensation is neither pleasant nor unpleasant in itself, i.e., it is non-hedonic. This desire is for the pain sensation *not to be occurring*. This desire is a mental state distinct from the pain sensation. This desire, in opposition to other mental states such as beliefs, has satisfaction conditions and a mind-to-world direction of fit. Broadly, a desire is satisfied once the world changes in such a way that it fits with the content of the desire. When the desire does not match the world, it is the world that ought to change. In a nutshell, according to desire theories a pain is unpleasant not because it *feels* unpleasant; rather, what it means to be unpleasant is that we desire not to have it. One of the important virtues of desire theories is that they can account for the heterogeneity of unpleasant experiences without claiming that there is a unitary feeling that all and only unpleasant experiences share.

The main problem for desire theories comes in the form of a Euthyphro dilemma. That is, there is a question with two possible answers and whereas one of these options is unavailable for desire theorists, the other is problematic to their own account. The dilemma is the following: do we desire a pain sensation not to occur because it is

unpleasant, or is the pain sensation unpleasant because we desire it not to occur? I will explain why this is a dilemma for desire theories and analyse what I consider to be the best possible solution that this approach can offer the dilemma. I will argue that the best way to deal with the dilemma is to argue that the property of being unpleasant is applied to the compound of the pain sensation and the desire, instead of such property only applying to the pain sensation. However, I will argue that the desire theories are unable to give a proper account for the desire that constitutes an unpleasant experience.

In Chapter Three I analyse the *distinctive feeling theory*. This view has a more intuitive and simpler answer to the Euthyphro dilemma. This theory accounts for the unpleasantness of pain by claiming that such unpleasantness is qualitative, a feeling, and a phenomenal aspect of experience. In other words, unpleasantness is a phenomenal property. According to the distinctive feeling view, we can explain that we desire a pain not to be occurring because it is unpleasant. However, this view has one important problem: the heterogeneity problem. The problem is that whereas this theory claims that there is a unitary feeling of unpleasantness in virtue of which unpleasant experiences qualify as such, there seems to be no unitary feeling of unpleasantness after careful introspection. I will show how the heterogeneity problem is fundamental for the distinctive feeling theory.

In Chapter Four, I analyse two theories that rely on the notion of dimension in order to account for the heterogeneity of unpleasantness. I call these the *dimensional theories of unpleasantness*. There is a way we could try to deal with the heterogeneity problem, while also thinking that unpleasantness is a phenomenal property. This solution relies on understanding the variability of unpleasantness in terms of the determinable-determinate distinction. This is a different attempt to deal with the heterogeneity problem, one that is different from that provided by desire theories. I think this solution is better than the one that desire theories can offer, since, in contrast to desire theories' solution, we can maintain the strong intuition that unpleasantness is something felt. In this view, being a pain entails being unpleasant. However, this solution has a fundamental problem: accounting for cases of non-unpleasant pains.

In Chapter Five I analyse potential cases of non-unpleasant pains. There is an intriguing issue regarding the existence of pains that are not unpleasant. In this chapter, I consider various cases that suggest that pain is not always unpleasant. I focus on two different types of cases that raise doubt about the idea that pains are always unpleasant: i) pain insensitivity and ii) pain indifference. If pain is the paradigmatic example of an unpleasant

experience, it seems odd to think that pain might not be unpleasant. However, various philosophers think that there are concrete cases where people do have pains that are not hedonic, where *pain asymbolia* seems to be the strongest candidate. In contrast to this current general consensus, I deny that there is strong enough evidence to conclude that people might experience pains that are not hedonic, i.e., experiences that are neither pleasant nor unpleasant.

In Chapter Six, I discuss another intriguing case that suggests the existence of non-unpleasant pains: masochism. Masochism might even suggest that pain is pleasant rather than unpleasant. I delineate different scenarios and reasons why people would seek out pain experiences. I categorise three different scenarios in which people might interact with pain: i) means-end masochism, ii) side effect pains, and iii) end-in-itself masochism. I propose that the unifying feature among cases of masochism is that people seek out pain. However, when people interact with an unpleasant pain experience sometimes they do it despite the unpleasantness of the pain, and sometimes they do it because of such unpleasantness. Even if pleasure might be involved in masochism, I conclude that there is no evidence of non-unpleasant pains.

In Chapter Seven I propose the hedonic determinable-determinate theory (HDDT). Even if there are no concrete cases of non-unpleasant pains, the discussion around such cases shows that we should have a theory of unpleasantness that allows the possibility of non-unpleasant pains. HDDT is also based on the determinable-determinate distinction and it accounts for the possibility of non-unpleasant pains. In this chapter I explain how the unpleasantness of pain should be understood based on the determinable-determinate distinction. Based on this distinction we can account for the variety of many different ways of being unpleasant. In a nutshell, I propose that unpleasantness should be understood as a phenomenal determinable property with different determinates u1, u2, u3, etc. I defend this approach and argue that it is better than any other theory from the previous chapters.

Before developing these ideas, there are two more things that I think should be mentioned: i) the motivation of this project, i.e., why I chose to write a thesis about pain and unpleasantness, why this subject is important in the philosophical debate, and ii) a preliminary understanding of the *explanadum*, i.e., I think that it will be useful to have an idea of the type of phenomena that are supposed to be explained. My initial interest came from wanting to understand sadness. More precisely, I wanted to understand why it felt bad to be sad or, in other words, what made this experience unpleasant. In order to try to

understand this unpleasant character, I started first by trying to engage with what we often refer to as 'physical pain', which seemed more accessible to study than the fuzzy and volatile 'emotional pain'. I now think that this distinction between physical and emotional is not very clear or useful. At the end, I think we can explain in a quite similar fashion in what consists for many different experiences to be unpleasant.

Another question that has particularly motivated me to examine the nature of pain and unpleasantness is what I discovered to be the *heterogeneity* of unpleasant experiences. That is to say, how to explain that many different unpleasant experiences are unpleasant but, at the same time, are so diverse and feel so different. To what could we appeal to explain the diversity within the unity of unpleasant experiences? It can be unpleasant to feel rejected by our colleagues, our friends, or our family. It can also be unpleasant when we cut or burn ourselves. It is unpleasant to be hungry or thirsty too. How can we explain what unifies all these divergent experiences? This question is key throughout this thesis, and I now have an answer for it.

The reasons to engage in understanding the nature of pain and unpleasantness are not only personal or biographical, of course. In recent years important attention has been given to pain and unpleasantness. Aydede's (2006) and Corns' (2017) compilations on the nature of pain are clear examples of the relevance of this subject. Moreover, the *Value of Suffering* project, which I was lucky to be part of, is another prominent example of the current interest in pain, unpleasantness, and affective experience in general, i.e., experiences that are pleasant or unpleasant. Moreover, pain and unpleasantness are very relevant in different aspects of our lives: when we go to the doctor, the hospital, the drugstore, etc. it is often because we are trying to find a way of getting rid of an unpleasant experience such as pain.

This connects with clarifying the *explananda* of this thesis. One of the main aims of this thesis is to explain what the unpleasantness of pain is in order to shed light on the nature of pain and unpleasantness. I think that we are relatively well acquainted with different unpleasant pain experiences such as the one that we get when we are burned, cut, generally injured, or when we have a headache. These are the kinds of experiences that I will shed light on. I account for the fact that such experiences are unpleasant, that they hurt. When you open and close your eyes there is some qualitative difference for you, there is something phenomenal that changes when your eyes are open that is different when your eyes are closed. However, the experiential changes that you have while opening and closing your eyes do not hurt. These visual experiences do not seem to be

unpleasant in themselves. In contrast, the sensation that you have when you pinch yourself is unpleasant. It is that unpleasantness that will be clarified in this thesis.

I focus on the unpleasantness of pain, although there are many more experiences that are unpleasant. There are other bodily experiences that are unpleasant even if they are not normally considered as a pain: hunger, thirst, nausea, itchiness, etc. There are other unpleasant experiences that have a bodily aspect, since they are felt somewhere in the guts or in the chest, that are emotional: grief, sadness, anxiety, disappointment, etc. Even if most of the discussion is centred on pain and its particular unpleasantness, the discussion inevitably touches upon many forms of unpleasant experiences. Moreover, by talking about unpleasantness, I also deal with many aspects of pleasure, or pleasantness. As you will see, many of the central authors that I discuss talk about pleasure and pleasant experiences rather that unpleasantness. I will show how these are mirror discussions to a large extent.

I provide a detailed analysis of the most prominent contemporary theories that try to account for the unpleasantness of pain. I explain which are the current most prominent candidates and point out their main difficulties. I consider various empirical cases and show that it seems that all pain experiences that people have are, in fact, unpleasant. I show that even if unpleasant experiences are bad in themselves, there are many different reasons why we might seek out unpleasant pain experiences. Finally, I propose my own account of unpleasantness, which is based on the determinable-determinate distinction. This theory can explain how unpleasantness is something felt while explaining how so many diverse unpleasant experiences all feel unpleasant. By doing all of this, this thesis will provide a detailed understanding of the nature of pain and unpleasantness.

CHAPTER 1: CONTENT THEORIES

1.0 Introduction

In this first chapter I will explain and analyse the theories according to which the notion of mental content can arguably account for the unpleasantness of pain. I will analyse the *content theories* of pain. The notion of mental content might seem very fruitful to elucidate many aspects of mental states, especially to give a naturalistic account of the phenomenology of experiences, so it is natural to begin this journey by exploring how such a notion can be used to explain pain and its unpleasantness. In this chapter I will tackle the main theories that account for pain and unpleasantness in terms of mental content. I will show, however, that content views face a fundamental problem. In order to do this, I have divided the chapter into four main sections.

The subject of the first section is *representational content* views. I start by clarifying how to understand what mental content is and how mental content is supposed to account for pain and its unpleasantness. I will focus on two representational accounts. I will refer to the first representational account as the *damage account of pain* (DAP). To put it simply, according to this approach the experience of an unpleasant pain consists in having a mental state that represents some form of bodily damage. I will show that this account is unsuccessful, and that the methodology lacks a very important feature. That is, it can't explain the unpleasantness of pain by appealing to such representation of bodily damage. The main problem is that the representation of damage, where damage is understood as bodily disturbances, cannot explain how unpleasant pains are motivational states. Given that DAP faces this difficulty, I move into exploring a second representational alternative.

The second view I will focus on is Bain's (2012) proposal. I will call this approach *Bain's account of pain* (BAP), also known as *evaluativism*. This is still a representational account but it adds an important feature. Namely, according to BAP an unpleasant pain consists in the experiential representation of a bodily disturbance as *being bad* for the subject. It is similar to DAP in the sense that an unpleasant pain is explained as a mental state that represents something, but according to this view in the case of an unpleasant pain this particular experience consists in representing something in an evaluative way. An unpleasant pain involves the representation of something as being bad for oneself. The fact that something is represented as being bad is meant to account for the fact that such experience is unpleasant. Hence this view is called *evaluativism*. BAP has important

problems, however. Among these, there is a salient difficulty: to explain how such representational content, even if it is evaluative, can in itself explain how an unpleasant pain motivates action. In order to tackle this, and other difficulties for the representational approach, I turn to different types of content theories that are meant to give a straightforward explanation for unpleasant pains being motivational.

The second section of this chapter discusses imperative content views, also referred to as *imperativism*. This is an alternative view that also explains unpleasant pains in terms of mental content. In contrast to the previous approach, however, this view explains the unpleasantness of pain as being constituted by imperative mental content, rather than purely representational content, as DAP and BAP propose. Here I am using "representational" to mean mental states with accuracy conditions, as opposed to satisfaction conditions. To put it simply, in the same way that we accept the difference between descriptive and imperative sentences, this approach considers that a similar distinction can be made for mental content. In contrast to the two representational accounts, in this case it is in virtue of the content of an unpleasant pain being imperative that we can account for unpleasant pains being motivational.

I will explain two different versions of imperativism. I will refer to the first one as *Martínez' view of pain* (MAP). This is defended by Martínez (2011, 2015). I propose that the strongest way of understanding MAP is that an unpleasant pain is a mental state constituted of a composite mental content, part of it is indicative and part imperative. The indicative part represents that you are having a bodily disturbance, and the imperative is ordering you to stop such disturbance. I will argue that this version is problematic since it is incapable of accounting for certain aspects of the unpleasantness of pain. Given these difficulties, I analyse a different imperative account. Klein (2007, 2012, 2015b) defends a different imperative approach, and I will refer to this version as *Klein's account of pain* (KAP).[1] According to this view an unpleasant pain is constituted of two different imperative commands: a command that orders you to maintain your body in a certain state and another command directed at the first command. However, KAP and MAP, as well as DAP and BAP, share the same difficulty: the messenger-shooting problem.

In the third section of the chapter I will focus on what I take to be the main problem for content theories in any of their versions. All content theories must confront the messenger-shooting problem. That is, beyond the particular difficulties that each type of content

[1] For a development of this theory regarding pain's intensity see Klein and Martínez (forthcoming), and for the imperative account in terms of pain signals see Martínez and Klein (2016).

theory has, one issue unites them all. These theories cannot offer an explanation of the non-instrumental badness of pain's unpleasantness. In other words, content theories cannot account for the fact that experiencing pain is bad in itself. These theories are able to explain various aspects of pain: they can explain, for example, why we perform certain actions when we experience pain. However, all content theories fail to clarify a crucial feature of an unpleasant pain by appealing to the content that is supposed to constitute these experiences: an unpleasant pain experience is bad in itself and that felt unpleasantness provides us with a good reason to act in order to stop experiencing it. It is because of this felt unpleasantness that we are justified in taking painkillers, for example. It is rational to take a painkiller because a pain is unpleasant. However, content theorists are committed to saying that it is non-rational to take painkillers, that a pain being unpleasant is not a good reason to take a painkiller.

Finally, in the fourth section, I will consider two solutions to the messenger-shooting objection. The first explains that it is rational to take painkillers based on instrumental reasons. The second accounts for the non-instrumental badness of pain by appealing to other mental states such as negative emotion or an aversion. I will show that the best strategy available to content theorists for dealing with the messenger-shooting problem is to appeal to aversion understood as a con-attitude, as desire for the pain not to occur. As a result, content theories are unable to account for the nature of an unpleasant pain by appealing to mental content; these theories have to renounce to the idea that content alone accounts for the unpleasantness of pain, and for the badness of such unpleasantness, in particular. Desire theories seem to be a better strategy for accounting for the badness of the unpleasantness of pain. I will explain this alternative account and its problems in Chapter Two.

1.1 Representational theories

The representationalist approach can be very useful for understanding various aspects of the mind. The discussion around representationalism is vast, but there are two central ideas that are worth pointing out: i) representational states are about something and ii) representational states have accuracy conditions. Let me explain what these two features mean before I move into explaining how we can try to understand an unpleasant pain experience in representational terms.

First, representations are *about* something. This is closely connected to the idea that mental states have content. In order to illustrate this, think of a thermometer. This mechanism represents something, i.e., it is about something; it is about the temperature in a certain place. Similarly, according to representationalism, we can understand certain mental states as representing something. Your beliefs, for instance, are about something. When you believe that you have a red apple in front of you, your belief is, roughly speaking, about a red apple. Other mental states can also be representational. We can also understand our sensory experience in this way. Vision, we could say, is about objects in the world. Vision informs us about the external world by representing it. You can have a visual experience about a red apple. This means that you are representing a red apple as being in front of you. Vision, in this way, is a way of representing objects. In other words, vision is a mode of representation.

Second, representations can be *accurate* or *inaccurate*. The thermometer may inaccurately indicate that it is 30 degrees Celsius even if the temperature is actually 35 degrees Celsius. Vision can also represent accurately or inaccurately. There are two main ways in which vision, and sensory experiences, may misrepresent. First, by misrepresenting a feature of an object; for example, you might have the experience of a red apple, but the apple is actually green; this is what is often referred to as an *illusion*. Second, by misrepresenting the presence of an object; for example you might have the experience of an apple, but there is no apple that is causing you to have such an experience; this is often understood as a *hallucination*.

It is also worth noting that there might be *veridical hallucinations*, i.e., a hallucination where what you are representing matches with an external object, but the causal chain between the experience of representation and the world is inadequate for the experience to count as a perceiving of the object. For example, you have the experience of a red apple in front of you because you took ayahuasca, a hallucinatory brew that induces you to have the experience of a red apple as if there was one in front of you. It happens to be the case that there is a red apple in front of you, but your experience of a red apple is not caused by that red apple. This would count as a type of veridical hallucination. Even if there is an important discussion about more details regarding these two aspects of representation, i.e., intentionality and accuracy conditions, I think that having this in mind will serve us to try to understand how we can account for unpleasant pains in representational terms.

1.1.1 Damage account of pain (DAP)

When we think of pain we often think of bodily damage. If we want to give an account of pain, it might be sensible to think that pain is strongly related to some form of bodily damage. When we cut, burn, or scrape our skin we often experience an unpleasant pain, for instance. Moreover, if we take into consideration some standard definition of pain such as the one provided by the International Association of Pain (IASP), there is a straightforward reference to damage. Pain, they say, is "an unpleasant sensory and emotional experience associated with *actual or potential tissue damage*, or described in terms of such damage" (my emphasis). It is unsurprising then that an initial representational account of pain will also refer to damage. Pain can be understood as the representation of bodily damage. However, before we enter into the details of this account, I propose that we need to be more explicit about the type of representational theory we will be dealing with.

We are dealing with a *strong version* of representational content account. In this strong version of representational content we assume two important features: i) the mental state that we are dealing with is purely representational, i.e., that the mental state's content constitutes all aspects of the phenomenal character of the mental state, and ii) all aspects of the mental state that we are trying to understand can be explained, clarified, illuminated, etc. merely in virtue of such content. This means, for example, that if we try to explain the phenomenal character of vision in terms of this sort of strong representational account, all aspects of what it is like for one to have a visual experience are representational, and that everything that there is to explain about vision's phenomenology can be clarified by appealing to the representational content that constitutes such visual experiences. The phenomenal character of experiences consists in their mental content.[2]

To illustrate this, let's consider what happens when we are having the experience of seeing a red apple. If we account for this according to the strong version of the representational approach, then everything there is to be illuminated about the phenomenology of this mental state is exhausted by its representational content. The phenomenology of the experience consists in having such representational content. This phenomenology is also sometimes referred to as the *what-it-is-like* of undergoing such visual experience.[3] The

[2] For a discussion on various aspects of representationalism in relation to phenomenal experience see Lycan (2015).

[3] See Nagel (1974).

subjective aspect of seeing a red apple is nothing beyond what that experience represents. That is, we should be able to understand what it is to have this sort of visual experience in reference to the experience's content, what the experience is made of, as it were. If we want to understand what it means to have an experience of seeing a red apple, we should understand what the content of this experience is. If we want to account for such content, we must account for what the experience represents. When we have the experience of seeing a red apple, there is a phenomenal property, which is an aspect of the visual experience and this consists in a specific representational content.

To put it simply, having a conscious qualitative experience of a red apple means that one is having a mental state representing an object as being red. Another way of trying to make sense of this is to put it in terms of properties or information; namely, having a visual experience of a red apple consists in representing an object as having certain properties such as being red, or representing certain information such as an object being red. There are, of course, many more details about the explicative power and complications of the representational approach in order to explain these and other mental states — Drestke (1995) offers what I take to be clear analysis of this type of representationalist view of the mind. However, given the subject matter of this chapter and thesis, I propose that we jump straight into analysing the strong representational approach for pain. I will refer to this strong version of representationalism simply as "representationalism" from now on.

How can we account for an unpleasant pain according to representationalism? What are unpleasant pain experiences representing? According to this view an unpleasant pain experience is representing bodily damage. Tye initially defended a view along these lines. He tell us that: i) "pains are sensory representations of bodily damage or disorder" (Tye, 1995, p.116), and ii) that in his view "pain experiences have a distinctive representational content and that this content is their phenomenal character" (Tye, 2006, p.99). Even if Tye's view on pain has evolved and changed from these two points, they are helpful to summarize the initial idea of explaining pain in terms of bodily damage.[4] Since Tye modified this view on the content of pain, I will refer to this approach not as his, but as a more general damage account of pain. This is the proposal:

[4] Tye has changed his view; he turned to defend a form of evaluativism (Cutter & Tye, 2014).

Damage account of pain (DAP)

An unpleasant pain is a mental experience constituted by the representation of part of one's own body as being damaged.

Let me explain what this proposal implies. First, it means that we can explain what an unpleasant pain is by referring to what this mental state is. An unpleasant pain is the representation of your own body as being damaged. To be in pain, and for pain to be felt the way it does, is to represent something as being a bodily part of yourself that is damaged, to receive information of your own body. If we want to explain why an unpleasant pain feels the way it does, for example, we should appeal to this content. In the same way that having an experience of something being red consists in representing an object as being red, to have an unpleasant pain is to represent something as being your own body and being damaged. This is what an experience of pain is supposed to be about, this is its content, and this content should capture the phenomenology of an unpleasant pain.

According to DAP, then, the content that constitutes an unpleasant pain may be accurate or inaccurate. Notice that this allows representationalism to elucidate cases such as the phantom limb phenomenon. "Almost everyone who has a limb amputated will experience a phantom limb — the vivid impression that the limb is not only still present, but in some cases, painful." (Ramachandran & Hirstein, 1998, p. 1603) People sometimes experience unpleasant pain as being in parts of their bodies, even if these bodily parts have been amputated. How can we explain that someone feels pain in a missing part of his or her body? DAP could explain phantom limb pain as a hallucination. You have a pain experience as being in your hand because you represent your hand as being damaged, but, since you have no hand, there is actually no damage in your hand. To put it simply, the fact that you represent a part of your body as being damaged does not imply that such a body part is actually damaged. A pain illusion, for example, would occur when you represent your hand as damaged and you do have a hand, but it is not really damaged. Your experience is still of a bodily damage of your own, but you are misrepresenting the object of experience. A pain hallucination, as I just said, could occur if you have that same representational content of your own hand as being damaged, but you do not have a hand.

Your experience is representing something, a part of your body as damaged, even if there is no such thing, since that part has been amputated.[5]

If this is correct, the representationalist could say that vision and pain are analogous. When you have a visual hallucination, there is something visual about your experience, but you are not really seeing things in the sense that your visual experience is misrepresenting. A visual hallucination is experienced as visual because what it means to have a visual experience is to represent an object as having certain properties such as colour, and the representational view can explain how, even if you do have a visual experience, the content of this experience might be inaccurate. Similarly, when someone has a phantom limb pain experience, there is something pain-like, in the sense that it feels like you are having an unpleasant pain, but this does not mean you are experiencing pain in the sense that you are accurately representing bodily damage. Pain hallucinations feel like they do by virtue of being representations of bodily damage, not because there is actual damage. You are really having a pain experience, i.e. you really are *in* pain, but the pain experience is hallucinatory. It does not count as a perceiving of a body part.

This gives a clearer understanding of how to understand what an unpleasant pain might be. However, this approach has some issues. Let me discuss two important ones, and propose that DAP is only able to give a solution to the former. The first issue is that the DAP proposal may seem insufficient to account for what constitutes an unpleasant pain. That is, we could have a mental experience with exactly the same content without the experience being an unpleasant pain. If the DAP definition were correct, this would mean that any mental state that represents your own body as being damaged should be an unpleasant pain. This, however, does not seem to be the case. There are other mental states that could represent such bodily damage without being an unpleasant pain. For example, you might have a visual experience with that same content; you might represent through vision that your body is damaged. For instance, you might see that you cut yourself. This, I take it, would not be an unpleasant pain experience. A visual representation of bodily damage is not an unpleasant pain. It does not hurt. If an unpleasant pain is nothing but the representation of your own bodily damage, then any representation of your own body as being damaged should be experienced as an unpleasant pain. This is not the case: some other mental states may also represent a part of your own body as damaged without constituting unpleasant pain experiences.

---

[5] One might also have veridical pain hallucinations. For example, one might represent one's own hand as being damaged, one's hand might be damaged, but the cause of one representing one's own hand as being damaged is not one's own hand being damaged.

This problem has a solution. The DAP defender could argue that an unpleasant pain, in contrast to other types of mental experiences, gives us a different *mode of presentation* of one's own bodily damage.[6] This mode of presentation is constitutive of the content of the experience, while the mode of presentation is the way the content is. So, strictly speaking, a visual representation of your own body damage is not constituted by exactly the same content as an unpleasant pain. This idea of mode of presentation is useful to explain how two different mental states can be about the same thing and yet be different experiences. For example, you might represent an object as having a given shape either through vision or touch. The content of these two experiences is about the same thing: both the visual and the tactile experience are representing the same shape, a cube. However, seeing or touching a cube are qualitatively very different types of experiences. One way to account for this difference is to say that these different senses can provide the same content with two distinct modes of presentation. They are representing an object as having the property of being a cube in a visual and a tactile presentation. Similarly, one could argue, when in pain we represent bodily damage in a different way than when we see that our body is damaged. This explains why pain feels different from vision; pain provides a different mode of presentation than vision.

However, there is another problem for DAP. It cannot explain why unpleasant pains are motivational in themselves, why we are motivated to perform certain actions by appealing to the content that is meant to constitute unpleasant pain experiences. There are many types of behaviours that an unpleasant pain itself motivates, i.e., many actions that we can render intelligible in virtue of an unpleasant pain. What motivates action is not some additional state, e.g., a desire not to be damaged, or a desire not to have an unpleasant pain, but the unpleasant pain itself. Let us focus for the moment on bodily-directed behaviour: you approach your finger to a candle flame, you feel a burning feeling, and so you quickly draw your hand back. If DAP is correct, then we should be able to explain why feeling a burn makes us act in virtue of the experience's representational content, in virtue of what an unpleasant pain ultimately is. We should be able to explain that an unpleasant pain makes us act in virtue of it being the representation of our own body as being damaged. However, the representation of our own bodily damage does not entail action. We cannot explain why we act merely in virtue of an unpleasant pain being the representation of our own bodily damage. Let me now explain why this is.

---

[6] The notion of mode of presentation could also be understood, I think, in terms of *experiential attitudes* (see Siegel, 2016, p. 15-18).

According to DAP, vision and pain are quite similar: they are mental experiences constituted by representational content and it is in virtue of their content that we can explain all of their features, e.g., why these experiences are qualitatively experienced they way they are. Vision represents information about objects: for example, it might represent the colour and shape of a red apple. However, representing colour or shape in itself does not entail that you will have any motivation to do something about these representations. Merely representing something that is either accurate or not, does not entail motivation for action. You might be motivated to do something about the apple, you may want to eat it if you are hungry, but this action is not explained solely in virtue of the representational features of your visual experience.

If this is correct, and pain is also explained based on the same representational approach, we still cannot explain why an unpleasant burn makes us withdraw from a flame. If an unpleasant pain is nothing but the representation of something that might be accurate or not, then there is nothing that explains why pain makes us act. The DAP supporter might argue that an unpleasant pain is motivational precisely because it represents *damage*. However, if according to DAP the representational content of an unpleasant pain is analogous to the content of a visual experience, except for the fact that they are different modes of presentation, then saying that the representation is about damage does not make any substantial difference. That is to say, if damage is representational, then it is nothing but the representation of a body part as being in certain conditions. Representing damage is, in this view, the representation of your own body as being in a certain way. Just representing your own body as being in a certain condition does not entail in itself that you act regarding such representation.

Let me illustrate this. You can look at your hand and have a visual experience of your hand. As you move your hand closer to the flame of a candle, you can see how you hand looks brighter. That is, you have a visual representation. This change can be understood as representing changing conditions about the light illuminating your hand in different ways as your hand gets closer to the source of light, as the flame becomes nearer. This representational content, which can be accurate or not, does not seem to be sufficient in itself to explain your actions, e.g., that you keep on moving your hand closer to the flame. We can explain why you continue moving your hand, but we cannot explain it by merely referring to the representational content that constitutes your visual experience. As an analogy, when you represent that your own finger is damaged, since this means that you represent that the condition of your skin is changing, it does not seem to be enough to account for your motivation to withdraw your hand from the flame. The fact that pain is

meant to represent damage as bodily changes is insufficient in itself to account for your withdrawal action. DAP is meant to explain what pain is and all aspects of pain, but DAP cannot account for this simple withdrawal action. The question left for the DAP supporter is: how to account for this motivational feature of an unpleasant pain while remaining a representationalist?

1.1.2 Bain's account of pain (BAP)

There is a quite straightforward solution for DAP. It consists in adding an evaluative component to the content that constitutes an unpleasant pain. It is in virtue of this content being evaluative that we are supposed to explain what makes the experience unpleasant and motivational. This approach is also referred as *evaluativism*. I take Bain (2012) to be the paradigmatic defender of evaluativism.[7] Here is his proposal:

> **Bain's account of pain (BAP)**
>
> A subject's being in an unpleasant pain consists in her (i) undergoing an experience (the pain) that represents a disturbance of a certain sort, and (ii) that same experience additionally representing the disturbance as *bad* for her in the bodily sense.

This evaluative twist is meant to explain, among other things, why unpleasant pains are unpleasant and thus motivate action. An unpleasant pain is not only representing *neutral* information that might be accurate or not. An unpleasant pain is constituted by the representation of *evaluative* information, an unpleasant pain conveys information about evaluative facts, i.e., an unpleasant pain represents objects as being bad. In this case, an unpleasant pain is the representation of a bodily disturbance that is represented as bad for oneself. Remember that in content theories we explain the various aspects of the phenomenology of experiences by appealing to their content. An unpleasant pain *feels unpleasant*, the fact that the bodily disturbance is represented as bad is meant to explain why the experience feels bad, and why it is motivational in virtue of feeling the way it does. An unpleasant pain is also meant to motivate because it feels unpleasant, so the evaluative content should explain why an unpleasant pain is motivational insofar as it is unpleasant. We do things when we feel pain in virtue of how an unpleasant pain feels, and

[7] Bain is not the first to propose a form of evaluativism. See Helm (2001, 2002), for example. However, I take Bain's to be the clearest and paradigmatic version of evaluativism.

this experience feels in the way it does in virtue of its content. After all, what it is about pain that motivates is, arguably, the way it feels, so if the content does not explain the motivation, it has not explained the feel. According to BAP, your motivation to act when you experience an unpleasant pain is explained in virtue of the evaluative aspect of the representational content.

We need now to say a few things about motivation and normativity. In order to better grasp how the represented badness of a bodily disturbance accounts for our motivation to act, we need to clarify what we mean by motivation. Given that we need to introduce a more careful understanding of motivation, it is a good moment to also say a few things about normativity. The distinction between these two things will also be important for the 1.3 section of this chapter, when I explain the messenger-shooting objection, which I take to be the main problem for content theories. The following is a good way of distinguishing between motivation and normativity:

> Motivating and normative reasons do have something in common in virtue of which they both count as reasons. For citing either would allow us to render an agent's actions intelligible. This is essential. For there is *a priori* connection between citing an agent's reason for acting in a certain way and making her acting in that way intelligible: that is, specifying what is to be said for acting in the way in question. In virtue of their differences, however, motivating and normative reasons make actions intelligible for quite different reasons … By contrast with normative reasons, then, which seem to be *truths* of the form 'It is required or desirable that I Φ', motivating reasons would seem to be *psychological states*, states that play a certain explanatory role in producing action. (Smith, 1995, p. 94-96)

One important thing to point out is that when we talk about motivation and normativity we often talk about motivational and normative *reasons*. In this sense, an unpleasant pain might be a motivation or a justification for acting. Let us focus first on motivating reason. A motivating reason can be understood as a *psychological state*, a mental state, in virtue of which an action is intelligible. In this way, having an unpleasant pain may be a motivating reason since we can explain an action, render the action intelligible, by referring to the unpleasant pain. For example, why did you withdraw you hand from the fire? The answer could be that it was because you had an unpleasant pain. More precisely, the unpleasant experience of being burned, which is a psychological state, renders your withdrawal action intelligible. The unpleasant pain itself is motivational because, the idea

goes, we explain the action by referring to this experience solely, without needing to explain the action in relation to other mental states such as desires. In more detail, according to BAP the unpleasantness of your burning experience, which is constituted by the representation of the bodily state of your hand as being bad, is what renders your withdrawing of your hand intelligible. The difference and advantage of BAP over DAP is the following: whereas DAP cannot render actions intelligible by merely referring to the representation of bodily damage understood as bodily disturbances, BAP can do so by referring to the representation of bodily disturbances as being bad for oneself.

This will become clearer if we take into account an evaluative judgment. Let us consider an evaluative judgment that appears particularly uncontroversial, such as that killing is bad. Suppose that you take killing to be bad or, in other words, you represent killing as *being bad*. If this is the case, we should be able to render some of your actions intelligible by referring to your psychological states, e.g. to your judgement of killing as being bad. According to motivational judgement internalism, there is a conceptual or necessary connection between making some evaluative judgements and being motivated to act accordingly. It is a classical conception about motivation that there is an entailment between certain mental states, motivation, and action. Davidson (1963) offers a paradigmatic description of this entailment:

> *R* is a primary reason why an agent performed the action *A* under the description *d* only if *R* consists of a pro attitude of the agent toward actions with a certain property, and a belief of the agent that *A*, under the description *d*, has that property. (Davidson, 1963, p. 687)

In other words, a primary reason is the combination of two mental states: a pro-attitude, such as a desire, and a belief. These reasons are explanatory since they render actions intelligible. If we take these explanatory mental states as motivating reasons, this means that experiencing such mental states entails motivation for action. Such mental states are motivating reasons, and by having such reasons we are motivated to act accordingly. For example, you might not kill someone, among many reasons, because you represent killing as being bad, you have such a belief, and because you desire to do what you think is right. Even if you are very angry with someone, you might still represent killing that person as bad, and you desire to act in accordance to your beliefs of what is bad. The combination of this belief and desire render your controlled behaviour intelligible. These mental states are

your reasons not to kill someone, which means that such mental states motivate you not to kill.[8]

Many other actions can be explained in virtue of your evaluative judgement. For instance, your vegetarian eating habits - the fact that you do not eat any animals - might be at least partially explained by referring to your judgement that killing is bad. In this case, the notion of killing is applied not only to people but also to sentient beings in general. The fact that you don't eat animals is a way of avoiding something that you take to be bad, that is, the killing of animals. There are, of course, many subtleties to how evaluative judgment motivates action. For example, the judgement alone might not be sufficient to be motivational, since it might need a relevant desire about wanting to not kill animals. Moreover, this is an oversimplification given that actions are often motivated by many reasons and in relation to competing reasons.[9] We might also think that what really motivates is not the desire and the belief themselves, but the *acknowledgement* that we have them. That is, we might think that a rational agent is necessarily motivated by the *recognition* of her reasons (Finlay, 2007). The necessary entailment between motivation and action is given by the recognition of reason, not by the reasons themselves.

However, beyond these and other subtleties about motivation, this should clarify how the evaluative twist of BAP is put forward to help us understand what makes an unpleasant pain motivational. That is, if we can render actions intelligible solely by appealing to the experience of an unpleasant pain, this means that such experience is motivational in virtue of its feeling. This feeling is explained by the fact that the bodily disturbance is represented as being bad, and this should explain why feeling an unpleasant burn is a motivating reason to withdraw our hand from a flame. However, there is a problem for accounting for motivation in terms of evaluative content, you might think: representing something in evaluative terms (i.e., as being good or bad) does not really entail motivation. In other words, an evaluative representation does not entail that it is a motivating reason, i.e., we might not be able to render an action intelligible by referring solely to such evaluative representation.

Let us illustrate this. Your judgment tells you that killing is bad in itself. You truly believe this and, as a consequence, you also think that we should avoid all forms of killing. This

---

[8] There is a debate about whether the motivational reason is the mental state itself, or *the fact* that one has such a mental state (see Alvarez, 2016). However, given that both views have limitations, I opt to stick to the standard version and take the mental states themselves to be motivating reason for actions.

[9] See Ruben (2009) for a discussion of competing reason or "con-reasons".

alone, however, might not be enough to explain any action. As Davidson's model proposes, we might also need a pro-attitude, such as a desire, in order to explain an action and, therefore, to account for the motivating reasons of an action. For instance, you might truly believe that killing animals is bad, but still have no motivation to do anything about it. If this is possible, you have an evaluative judgment that does not serve to render any action intelligible, the evaluative representation in itself does not entail motivation. If this can be the case, then it is problematic for BAP. This is because according to BAP, unpleasant pain's phenomenology is constituted by an evaluative representation, and if this feeling is in itself motivational, then the evaluative representation should be a motivating reason in itself. However, as it happens with evaluative judgment, it does not seem that evaluative representations entail motivation because we cannot explain action by referring solely to such evaluative representations. A mental state might be constituted by a representation of something as being bad, and still not be an explanation for action.

I can see two ways in which BAP can offer a solution to this. First, a BAP supporter could argue that the connection between an evaluative representation and motivation is *defeasible*. That is to say, that you will be motivated to act if you have a reason only if you are rational. For example, Smith (1995) proposes this kind of model in relation to reason:

> [T]here is an analytic connection between the desirability of an agent's acting in a certain way in certain circumstances and her having a desire to act in that way in those circumstances if she were fully rational… If claims about what we have reason to do are equivalent to, or are in some way entailed by, claims about what it is desirable for us to do… then it follows that there is a plausible analytic connection between what we have reason to do in certain circumstances and what we would desire to do in those circumstances if we were *fully rational*. (Smith, 1995 p. 109, my emphasis)

That is to say, this means that an evaluative judgment would be a reason for action if our fully rational selves would have a desire to act in accordance to such evaluative judgement. If this happened, then the combination of such judgement plus the desire to act in accordance to the judgment would constitute a reason for action. However, this connection between rationality and desirability will not be useful for BAP in order to account for motivational reasons. This is so for two main reasons. First, Smith is giving criteria for *normative reason*. That is, the element of rationality is meant to explain when an evaluative judgment is desirable, i.e., when such judgment entails a normative reason. For the moment we are only concerned with an unpleasant pain being a motivating reason.

Second, even if Smith's requirement of desirability is about normative reason, we might argue that the evaluative judgment alone is not sufficient to give us a reason for action, since it also needs a desire for acting in accordance with the evaluative judgement in order for these mental states to count as a normative reason for action. Going back to the previous example, your judgement about killing animals as being bad is not a normative reason for action in itself, not even when you are fully rational. You need also a desire to act in accordance to that judgement, for it to be a normative reason not to kill animals.

What happens when we try to argue something similar for unpleasant pain as being a motivating reason, in the light of BAP? This strategy is not going to work. That is, we could try to argue that an unpleasant pain is only a reason for action if we are rational. This appears strange, however. This is because we are confusing motivational reason and normative reason. Smith claims that a normative reason requires some form of rationality. In contrast, being motivated when one feels an unpleasant pain does not seem to be the kind of situation where one has to be rational. This is because the unpleasant feeling of pain in itself seems to motivate action, beyond the fact that it also justifies action. However, BAP does not seem to capture this motivational feature of an unpleasant pain by claiming that the content of such pain is evaluative. Discussing rationality does not serve BAP when trying to explain how an unpleasant pain is a motivating reason. In the same say that an evaluative judgment in itself does not seem to entail motivation, the evaluative content of pain in itself does not seem to entail motivation either. However, such content should entail motivation if we accept that the unpleasant feeling of pain is in itself motivational, i.e., that we can render actions intelligible in virtue of this unpleasant experience. This is a major problem for BAP.

There are additional problems for evaluativism. The first one is that BAP might be too demanding if we are to account for an unpleasant pain. Some may mistakenly think that the BAP proposal is that pain's unpleasantness consists in an evaluative judgment, which requires sophisticated mental states that not all creatures that experience unpleasant pains have. A BAP supporter has a straightforward answer for this: pain does not consist in an evaluative judgment. An unpleasant pain is a *perceptual experience*; it is a perception of evaluative states. In the same way that we represent killing as being bad where badness is predicated of killing, when we experience an unpleasant pain, there is a bodily disturbance that is represented as being bad, where badness is predicated of the bodily disturbance. We might worry, nonetheless, that perceiving something as being bad is much more sophisticated than perceiving other qualities of objects such as colour, size, location, etc. One might worry that whereas human babies and many other animals experience

unpleasant pains, they might not be equipped to perceive something as being bad. Unless we explain in more detail what it means to have such evaluative perceptions, BAP looks less promising that we might have initially thought.

The second issue is related to this and is mainly an ontological worry. It is not clear what it means for some object of the world to be bad and, more specifically, what would count as badness when it comes to our body parts. Let us recall that this is a representational theory and that this means that the represented content can be accurate or inaccurate. The content of an unpleasant pain experience is informing on a bodily state as being bad. What does this imply when the content is indeed accurate? To put it simply, in the same way that we can represent objects as having certain features like colour, shape, or size, and this content can be accurate when the object has in fact the represented colour, shape, or size, there must be a situation where the bodily disturbance represented as being bad is actually bad. But what would this badness be? The representational view is often used to give a naturalistic explanation of the mind. So BAP could propose a naturalistic explanation of bodily badness such as an impediment of proper functioning of the body, lack of homeostasis, etc. However, I think that the problem runs deeper. That is, it is not obvious that we can provide a naturalistic account of normative properties such as being bad.

A good way of trying to make sense of this problem is to refer to Moore's (1903) open question argument. Moore argued that we couldn't give a definition of normative concepts by referring to purely naturalistic phenomena that are not normative themselves. BAP supporters are susceptible to this critique if they offer a purely naturalistic account for the badness of pain. Bodily badness, in so far as it is evaluative, is a normative concept. The critique is that each time that we try to define something as normative, such as right or wrong, or good or bad, in terms of something non-normative, such as purely natural phenomena, this always leads to an open question. This open question shows that we have not ultimately explained what makes something normative. Every time that we give a non-normative description of the normative feature that we are trying to explain, we can ask an open question about the non-normative phenomena, which is meant to show that we have not really given an account of the normative concept. Smith (1995) explains a similar worry that Ayer (1936) pointed out:

> Ayer's objection to definitional naturalism, following Moore, is that since for any natural property F it is possible to think that x has F while at the same time thinking that x is not right, it follows that, for no natural property F is it self-contradictory to claim that x has F but x is not right. But it would have to be

> self-contradictory, if it were possible to give a naturalistic definition of rightness. Therefore, it is not possible to give a naturalistic definition of rightness. (Smith, 1995, p. 36)

Something similar can be applied to BAP. If we try to account for the badness of an unpleasant pain in purely naturalistic terms, such as some detailed bodily disturbances, we can still think that such disturbances are not bad. Given that this is the definition of an unpleasant pain, it should be self-contradictory to think that some given bodily disturbances are not bad. But it is not self-contradictory and, therefore, a naturalistic account of the badness of an unpleasant pain cannot ultimately be a proper account. That is to say, if we ask what kind of phenomenon in the world - something that can be objectively measured - accounts for the badness of an unpleasant pain, we will find out that we cannot provide a satisfactory naturalistic explanation of such badness. If the purpose of BAP is to provide a representational account that explains the unpleasantness of pain in a similar way to how we could account for vision (and other perceptual experiences) by merely referring to the representation of information about natural phenomena, then BAP leaves an open question: are such and such bodily disturbances really bad? This, the idea goes, should appear as a self contradictory question and given that it does not look so, this shows that we haven't really provided an explanation of what the badness of an unpleasant pain is.

If we accept Moore's open question argument, then DAP and BAP have a quite similar problem. Namely, I take it that an important aim of representational accounts is to provide a similar explanation across sensory modalities. The phenomenology of vision is a good example of how a representationalist account helps us to understand this phenomenon in terms of the representation of natural properties. We can explain various features about what it means to have visual experiences in terms of the representation of light bouncing off the surface of objects, for example. The phenomenology of visual experiences can be explained to a large extent in terms of the representation of natural properties. This is one of the main virtues of the representational account for understanding phenomenology: that it serves to develop a naturalistic approach of phenomenology. If we want to provide an ultimately naturalistic explanation of the unpleasantness of pain, it is unclear how the representation of such purely naturalistic properties is going to shed light on the unpleasantness of pain. In other words, it is hard to see how normative properties, such as being bad, are motivational in themselves, if they are ultimately the representation of naturalistic properties.

In the case of DAP, once we understand bodily damage in terms of some precise bodily disturbances, then such disturbances do not seem to elucidate why unpleasant pains are motivational. BAP can try to explain such a motivational aspect by appealing to the represented badness that constitutes an unpleasant pain, but then once we try to make sense of such badness in naturalistic terms, it seems that we fail to properly explain what that badness is. As we have seen, the representational account does not seem to be very helpful in shedding light on how an unpleasant pain is motivational in and of itself.

In the next section I will analyse different types of content theories that set out to give a straightforward explanation for the motivational aspect of an unpleasant pain. These are the imperativist content theories. However, once I have explained this approach, I will tackle what I take to be a unifying problem for representational and imperative content theories. In section 1.3, I will discuss the messenger-shooting problem, which shows that none of the content theories can explain how an unpleasant pain is a normative reason for action.

1.2 Imperative theories

According to this approach the content of the unpleasantness of pain is imperative. In the same way that we accept the difference between indicative and imperative sentences, imperativists claim that this distinction can be applied to mental content. There are mental states that have an indicative content, such as visual experiences. These mental states are similar to descriptive sentences insofar as their function is to report and describe how the world is, and this information might be accurate or not. In contrast, mental states such as unpleasant pains are similar to imperative sentences, since they command us, they tell us to do something. Whereas the representationalist views of content tried to explain the unpleasantness of pain merely in indicative terms, by telling what unpleasant pain experiences represent, imperativist views account for the unpleasantness of pain in terms of what the experiences command us to do.

A good way of grasping the difference between these two types of mental content is to appeal to the notion of *direction of fit* (Anscombe, 1957; Searle, 1979). Mental states such as beliefs and visual experiences have a mind-to-world direction of fit. They represent the world accurately or inaccurately. They are accurate when “the mind fits the world”, that is, when the content of these mental states matches with the way the world actually is. The mind ought to match the world. For example, if you believed that there isn’t a red apple in

front of you, but there is in fact one, then it is your belief that ought to change in order to fit the world and to be accurate. In contrast, mental states such as desires have the opposite direction of fit. These mental states have a world-to-mind direction of fit. These mental states are not susceptible to being accurate or not. Instead, these other types of mental states can be satisfied or unsatisfied. They are satisfied when "the world fits the mind". When the content of a desire does not match with how the world is, then it is the world that ought to change for the desire to be satisfied. The world ought to match the mind. Roughly, your desire to eat a red apple would be satisfied once you actually eat a red apple. According to the imperative views, unpleasant pains have a desire-like direction of fit.

But what is the benefit of explaining pain in terms of an imperative content? A primary aspect of the imperative approach is to account for the motivational force of certain mental experiences. In contrast to the representational account above, the imperative content is meant to capture the motivational force of an unpleasant pain. According to imperativism, many of our bodily experiences such as hunger, thirst, itches, and pains are motivational because they are constituted by an imperative content (Hall, 2008). These experiences are motivational. If feeling hunger in itself gives us a motivating reason for action, then we should be able to explain this action by appealing to the imperative content that constitutes such experience. Desire, for example, is a typically motivational mental state. This is because, we can argue, the proper function of a desire is that the world ought to change so that it fits with the content of the desire. This is why we can render actions intelligible by appealing to a desire. If you desire a red apple, then you may act in order to get the apple. If we take hunger to be a motivational reason, it is because we can explain behaviour by referring to such experience, and we can explain that it is motivational because it has the same direction of fit that desires have. Hunger is a motivating reason in virtue of being constituted by an imperative content. In the same way that one could think that desires are inherently motivational, given that they have a world-to-mind direction of fit, if unpleasant pains are conceived as having a similar direction of fit, then pains are also inherently motivational. This, I think, gives us a general idea of what the imperativist accounts have in mind, so we can now take a look into what they propose more specifically.[10]

[10] See Bain (2011) for an explanation and critique of imperativism.

1.2.1 Martínez' account of pain (MAP)

What is, exactly, the content of an unpleasant pain, when it comes to imperative content? What is such unpleasantness telling us to do? If imperativism is supposed to offer an explanation of many bodily experiences such as hunger, thirst, itches and pains, what are the commonalities and differences among these experiences' content? Broadly speaking, the commonality is that they all have imperative content; they all tell us to do something and this is how imperativists account for these experiences being motivational. The difference is that a different type of command constitutes the content of these diverse motivational experiences. Roughly, imperativists think that hunger is constituted by the command 'Eat!', hunger by 'Drink!', itch by 'Scratch!' and so on. With all this in place, we should try to explain what precisely the content of an unpleasant pain is.

I think that a good way to start with imperative content is Martínez' proposals (2011, 2015). Martínez (2015) tells us that a pain is constituted by the content: "See to it that bodily damage *d* does not exist." (2015, p. 2261) I think, nonetheless, that we need to make a modification to this proposal in order to make it stronger and to analyse it in its best shape. There are two main problems with this definition. The first one is that it includes the notion of damage, which is problematic if damage implies some form of normativity, something that is bad. This would collapse into a form of evaluativism, and we should avoid including evaluative terms in an imperative account of pain, since these are meant to be contrasting theories. This is not problematic if we understand damage as a representation of bodily disturbances. However, as we saw in the first version of representationalism, this representation seems to be inert, and it does not shed light on how pain is a motivating reason in itself, so we need to add a motivational component.

The second and more important problem with Martínez' definition is that it does not *explicitly* differentiate the sensory aspect of an unpleasant pain from the hedonic one. That is to say, the sensory aspect refers to the phenomenology of the experience in virtue of which it is comparable and distinguishable from other sensory experiences such as vision, touch, etc. This aspect of the experience is not meant to explain how an unpleasant pain is a motivational reason, how such experience might account for actions. In contrast, the hedonic aspect of an unpleasant pain, its felt unpleasantness, is responsible for unpleasant pains having a motivational dimension. Most of the theories of the nature of pain and unpleasantness distinguish these two aspects. One of the main reasons for this is to be able to deal with cases where it seems as if people have pain experiences that are not unpleasant —I will go into the details of this in Chapter Five and Chapter Six. Here is

what I take to be a better formulation of Martínez' proposal, without changing the essence of his view:

> **Martínez' account of pain (MAP)**
>
> An unpleasant pain is constituted of i) an indicative content "There is a bodily disturbance *d* in your own body", which accounts for the experience being a pain, and ii) a hedonic aspect about the indicative content "See to it that bodily disturbance *d* does not exist!", which account for the unpleasantness.

This formulation has several advantages. There are two aspects of this proposal that I think are particularly relevant. First, by clearly distinguishing the two aspects of an unpleasant pain, it provides a straightforward explanation of what constitutes the motivational character of an unpleasant pain. An unpleasant pain is a motivating reason in virtue of its unpleasantness. This sheds light on the intuition that unpleasant pains are motivating in virtue of being unpleasant. If unpleasantness is what makes an unpleasant pain a motivating reason, we can explain this in terms of such unpleasantness being constituted by an imperative content. Given that the unpleasantness of pain is imperative, and that imperatives are desire-like states, we can render actions intelligible by appealing to the imperative mental content that constitutes an unpleasant pain. The fact that an unpleasant pain is partially constituted by a command explains why having an unpleasant pain is a motivating reason.

Second, this formulation of MAP maintains one of Martínez' main purposes, i.e., to capture that the unpleasantness of pain does not prescribe a particular action. The unpleasantness of pain gives us a motivating reason for action, but there are many different actions that could be adequate in order to deal with a pain experience. When we get burnt, we withdraw from the source of heat —since we often get burned in this way— but we might also stop moving when we twist an ankle. These are two different types of action in response to different unpleasant pains. Moreover, we might behave in different ways in response to the same pain experience; sometimes when we feel a burn, we withdraw, but we might also put the burnt area of our bodies under cold water. The key point is that these different actions can all be explained by appealing to the same content that constitutes the unpleasantness of pain, "See to it that bodily disturbance *d* does not exist!". An imperative content implies that the unpleasantness of pain has a world-to-mind direction of fit and, just as with desires, this is meant to constitute a motivating reason for action. This sheds light on how unpleasant pains are motivational because they are unpleasant.

However, MAP also has to face some difficulties. Let me tackle a few of these. The first one is that it is not clear how MAP could account for the variations between how different pain experiences are felt. That is, the content of an unpleasant pain is meant to capture the phenomenology of the experience, yet MAP is proposing a single content to account for unpleasant pains. This would imply, then, that there is only one way in which unpleasant pains feel, which is, I would say, obviously false since the pain produced by a cut and a burn feel quite different. This is even more salient when we consider a headache, a stomach-ache, etc. There is a solution for this. A MAP defender could argue that this proposal is an outline; that we would have to be more specific about the precise type of content that would constitute each specific unpleasant pain. The variability in terms how experiences are felt can be accounted for by the more precise indicative content, for example. I think this would be the most natural way for MAP to deal with the phenomenal variation. We could explain the felt variation among different unpleasant pains in terms of the differences in the indicative content that constitutes the experiences. This, however, is still problematic. This assumes that there is no variation regarding the unpleasantness of pain. It seems that there is also variation in relation to the unpleasantness of pain: variations in the way the unpleasantness of pain is felt. The different ways of being unpleasant is something that MAP cannot clearly capture. I will go back to this issue towards the end of this chapter, and in much more detail once I talk about the heterogeneity problem in the next chapters.

Second, there are some unpleasant pains with a much less intuitive explanation, if we follow MAP. Headaches are a good example of this. That is, if we accept that at least some headaches are the result of dehydration, it seems odd that the content of a headache orders the elimination of a disturbance in your head, when the cause of a headache has nothing to do with something in your head. One of the advantages of MAP is that it is meant to explain the kind of behaviour that we normally perform when we experience pain. For example, if you cut your finger and experience pain, you do something with you finger. If you twist your ankle, you do something with your ankle. However, this kind of story does not seem to apply to headaches, which are a very common type of pain. A MAP supporter could say even if dehydration headaches are caused by lack of water in our bodies, their unpleasantness is actually constituted by a command of the type “See to it that such and such bodily disturbance *in your head* does not exist!” In the face of it, I think that a MAP supporter has two alternatives. Either: i) accept that headaches are a type of pain experience that is systematically telling us to do something completely useless or, if this first option sounds too strange, ii) to conclude that headaches are a type of pain but

their unpleasantness must be explained in a different fashion. I don't think either of these options implies a catastrophe for MAP, but they do show make the view less attractive.

A third possible problem for MAP, which I believe to be much more substantive, comes in relation to the imperative content being sufficient to account for unpleasant pains being a motivating reason for action. MAP is supposed to explain what constitutes the motivational dimension of an unpleasant pain, but such imperative content might not be sufficient to count as a motivating reason for action. The idea is that we are not motivated to act merely in virtue of receiving a command: being told to do something in itself does not count as a motivating reason to act. For example, imagine that you are at a dinner party and the young child of one of your friends comes and tells you all "It is late, go to bed!" (Bain, 2011) This, I take it, is not obviously a motivating reason to act. You might be completely unmotivated to follow this order. If it is possible to remain completely unmoved by this order, this shows that a command, in itself, does not entail a motivating reason for action. This might not even give you a motivating reason that is not strong enough to override other competing reason. It is not as if you would do what this child tells you if you didn't have other stronger motivations. You are really completely unmotivated after receiving this order, even if it is a command. If the police came and told you to stop the party, on the contrary, you might probably have a motivating reason to stop it and go to bed. How could we account for this difference? One way to account for this and other aspects of an unpleasant pain is to turn to the next imperativist account.

1.2.2 Klein's account of pain (KAP)

Klein (2015b) proposes a different account of pain and unpleasantness. There are two aspects of his account that can be helpful to explain what else we need in order for a command to be a motivating reason. That is, the command has to i) come from a source of authority, in this case it is our own body, and ii) we must care about the command, since the command comes from a source of authority. These two elements are quite intertwined in Klein's account. He proposes that the commands that constitute a pain are issued by our own bodies, that these commands are in the best interest for ourselves, that we treat the body as a practical authority, and given that we care about the issued commands, this explains why such commands constitute a motivating reason for action:

> Commands that motivate are those that are issued by a source that we accept as having the authority to direct our actions, and so whose commands give us

> certain reasons to act… By analogy, the body is a bit like the state. The state issues imperatives in the form of laws and other directives. We are like citizens who accept the authority of the state, and so we are moved by what it commands... [However], the body is more like a dim-witted, paranoid king: well-intentioned but error-prone and obsessed with self-preservation, and whose commands coincide only roughly with the flourishing of his subjects. (Klein, 2015, p. 72-74)

Going back to the previous counterexample for MAP, we could explain that the difference between receiving an order from a child and from a police officer is that the latter has an authority that the former lacks, and so we care about the officer's command in a way that we don't about the child's. When it comes to pain, our bodily commands are rather like the command of the police. Let us have a look into Klein's proposal so as to unpack it, and show some of its strengths and weaknesses:

> **Klein's account of pain (KAP)**
>
> An unpleasant pain is constituted by two imperatives: i) a command "Keep B from E (with priority P)!", which accounts for the experience being a pain sensation, and ii) another command, "Don't have pain sensation!", which constitutes the unpleasantness of the experience.

This content is meant to account for all aspects of the phenomenology of an unpleasant pain. According to KAP, we can account for how different unpleasant pain experiences feel by referring to variations in the content of the first imperative. There are three main elements that constitute such pain feeling: "*B* stands for a particular body part, *E* a nominalised passive gerund phrase, and *P* a ranking function" (Klein, 2015, p. 57). In summary, *B* is meant to capture the felt location of a pain experience, a pain experience is about a body part and this determines where the pain is felt. *E* is a commanded action that tells us how to treat the bodily location; this command also constitutes how the experience feels, i.e., different ordered actions account for different felt experiences. *P* is meant to capture the intensity with which the pain is felt. According to KAP, we have a list of ranked mental states that serve us a guide for action, a sort of mental to-do-list; the higher the command is in the ranking, the more intense the experience is. For example, when you feel pain because you twisted an ankle, this experience consists in a command "Keep your ankle from bearing weight!", where this command has some priority in relation to other motivational mental states.

All these three elements, it is worth insisting, are only to account for the sensory aspect of a pain. This sensory aspect, according to KAP, is motivational even if it is not unpleasant in itself. The unpleasantness of the experience is explained in virtue of a second-order command, a command about another command. Particularly, the unpleasantness of pain consists in a command to stop a specific pain sensation. That is to say, the unpleasantness of pain in this view consists in a command telling us to stop having the first order command that accounts for the experience being a pain sensation. This is an important difference between MAP and KAP. Whereas MAP took the pain aspect of the experience to have indicative content, KAP takes that same aspect of the experience to have imperative content. If one wants to be imperative purists, KAP offers an option where all the content of an unpleasant pain is constituted by imperative content.

Beyond this and other possible advantages that KAP might have, there is a major problem for imperativists view. I take this to be the main problem for all versions of content theories. This should show that beyond the nuances of the type of content that constitutes an unpleasant pain, there is a feature of felt unpleasantness that none of the content theories can capture: the fact that unpleasantness is non-instrumentally bad.

1.3 The shooting-messenger problem

The messenger-shooting problem is that having an unpleasant pain is something bad in itself, but content theories cannot explain this feature by appealing to the content that is meant to constitute the unpleasantness of the experience. But first, what does it mean for an experience to be bad in itself? We could try to understand this in two ways. In a negative formulation it means that you are better off by not having an unpleasant experience, all else being equal. In a positive formulation it means that if you were having an unpleasant experience, you would be better off if you didn't have it, all else being equal. The badness of an unpleasant pain seems to be intrinsic. Moreover, if unpleasantness is something phenomenal, something that we feel, then this badness is also phenomenal. Unpleasant pains feel intrinsically bad. When you have an unpleasant headache, an unpleasant hurtful burn experience, an unpleasant cut, etc., there is something about those experiences that feels bad. Everything else being equal, you would have been better off without experiencing such unpleasant experiences. This is a crucial feature of unpleasant pains and other unpleasant experiences, but we cannot capture this feature by appealing to the content that different content theories propose.

The badness of unpleasantness is very important in relation to motivation and normativity. Let us now focus on the latter. It is in virtue of this felt badness, i.e., in virtue of unpleasantness, that we have good reasons for getting rid of our unpleasant pains. It is *desirable* to get rid of our unpleasant pains because they feel bad. If we have good reason for an action, this means that we are justified to perform such action. The fact that it is desirable for us to get rid of an unpleasant pain that feels bad, shows that the unpleasantness of pains provides us with normative reason. When you have a headache, it feels bad, and since it is feels bad, this provides you with a good reason to take a painkiller. Taking painkillers is, therefore, rational. However, content theories are committed to saying that taking painkillers is actually non-rational. To make this clear, let us go back to the distinction between motivating and normative reason. Whereas the former can be understood as an internal mental state in virtue of which an action is intelligible, the latter is something in virtue of which *it is desirable* or required that you perform such action. An unpleasant pain is a motivating reason because we can render actions intelligible in virtue of such mental states, and an unpleasant pain is also a normative reason because it is desirable that we act in order to avoid having such an experience.

Let us picture the next scenario to make this clearer: Othello believes that Desdemona has betrayed him very badly. Othello believes that Desdemona has wronged him so badly that he has to kill her, but in fact it was really Iago who betrayed Othello. In this case, Othello has a motivating reason to kill Desdemona. Othello's belief, together with his desire, constitutes a motivating reason. We could explain Othello's behaviour by appealing to these mental states. However, these mental states do not imply that Othello has a normative reason, a good reason, to kill Desdemona. In other words, it is not desirable or required, given the circumstances, for Othello to kill Desdemona. Another way of putting this, following Smith (1995), is that if Othello were fully rational and knew all of the relevant facts, he would not advise his actual self to kill Desdemona. It is not desirable for Othello to act in accordance to his own belief. Othello has no *good reason* to kill Desdemona, because Desdemona did not betray him. Now, when it comes to an unpleasant pain, the unpleasantness of the experience is a motivating reason and also a normative reason. Since an unpleasant pain feels unpleasant, and unpleasantness is intrinsically bad, we have a good reason to get rid of the unpleasant pain. An unpleasant pain, in virtue of being bad in itself, is a normative reason.

Let us begin to explain how representationalists cannot account for an unpleasant pain being a normative reason. I will only focus on BAP, since it is a development of DAP, the

initial proposal based on the representation of damage. First, given that the representational content can be inaccurate, representing something as being bad does not imply that we have good reason for action. It is not necessarily desirable to act in accordance of a representation of a bodily part represented as being bad when the content is inaccurate, i.e., when the bodily disturbance is not really bad. We would have a good reason to act, you might think, only if our representations were accurate: if you represented with absolute certainty that something bad is going on in your body, then this would give you a good reason to act. You would have a good reason, it would be desirable, to do something about the body part where there is something bad occurring. However, this is not the nature of representations: representations might be inaccurate, they may misrepresent. For example, you might represent an inoffensive caress as being bad. Even if you represented this caress as bad, that does not mean that you have a good reason to stop the caress. The caress is, in fact, inoffensive. In the same way that Othello might be wrong about Desdemona betraying him, the content that constitutes our unpleasant pain experiences as being bad might be inaccurate. In the same way that it is not desirable for Othello to kill Desdemona, it is not really desirable to stop an inoffensive caress.

But the problem with BAP runs even deeper. BAP cannot account for an unpleasant pain being a normative reason for experience-directed actions, not even when their content is accurate. BAP can account for an unpleasant pain being a normative reason, a good reason, for *bodily-directed* actions, such as moving away from a source of damage, when the content is accurate. That is to say, when you represent that something bad is happening to your body because it is burning, then it is indeed desirable that you act in order to stop being burned if you are really being burned. You are justified to perform these actions in virtue of your unpleasant pain's representational content being accurate, because what you represent as being bad really is bad. It is rational, then, to perform body-directed actions when your pain is accurately representing something bad taking place in your own body.

However, BAP cannot account for *experience-directed* behaviour, such as taking painkillers, not even when the content is accurate. For BAP, an unpleasant pain can be undesirable because of the badness that it is representing, but it cannot be undesirable also in virtue of representing badness. This might seem like a subtle difference, but it is very important. As I said, felt unpleasantness is taken to be bad in itself. When you have an unpleasant pain, the bare fact of having this experience is bad, and gives you a normative reason to act. You would be better off if you didn't have it, all else being equal. This feature cannot be explained by referring to the representation of badness in a bodily part,

not even when the content is accurate. According to BAP, we cannot explain that it is bad to represent badness. But if this were right, then we cannot explain why it is rational to take a painkiller. For BAP there is no reason why you should take a painkiller, because simply representing something as being bad is not in itself bad. BAP is not capturing a crucial feature of an unpleasant pain. That is, we cannot explain in BAP terms why we have good reasons to stop experiencing an unpleasant pain.

Jacobson (2013) explains the messenger-shooting problem with an analogy from where the objection gets it name. When a messenger brings bad news to the king, it is non-rational for the king to kill the messenger as a way of dealing with the bad news. Receiving the message is not in itself bad, what is bad is what the message informs the king about, assuming that the bad news is accurate. If the messenger informs that people are dying in the kingdom, the message isn't bad: it is bad that people are dying. Whereas it would be rational to do something to stop people from dying, it would be non-rational to shoot a messenger who brings bad news as a way of dealing with the content of the news. It is not desirable to shoot the messenger. The king would have a good reason to send a doctor to the kingdom to stop people from dying; this action directed at the content of the news would be desirable. In contrast, the king has no good reason to shoot the messenger, since the fact that someone gives bad news is not in itself bad.

BAP is analogous to the messenger situation. Simply representing badness in your body is not bad in itself, as it is not bad in itself to receive a message about something bad. What is susceptible to being bad is what is happening in your body, in the one case, and what is bad is that people are dying in the kingdom, in the other. In the same way that it would be non-rational to shoot the messenger, i.e., it is not desirable or required to kill the messenger, BAP is committed to say that it would be non-rational to stop representing bodily badness. However, the theory should explain why it is actually rational to do something in order to stop experiencing an unpleasant pain. When we take a painkiller that is precisely what we do: we stop the experience. If it is rational to take a painkiller, but BAP is committed to say that it is non-rational, we must conclude that BAP is wrong.

Imperativism faces very similar difficulties to representational accounts regarding justification for action. First, imperativism has difficulties in justifying body-directed behaviour. Receiving a command does not entail that we have a good reason to obey and act in accordance with such an order: there are commands that we should not follow. The fact that we receive a command to do something about a bodily disturbance does not entail that we have a good reason to follow such an order. Suppose now that you receive a

harmless caress and, for some reason, you receive a body command to stop such a disturbance. Would you have a good reason to stop the caress merely in virtue of receiving this bodily command? No, you wouldn't, a harmless caress is not doing anything bad to you. The imperativist could argue that even if the commands are not always useful, they often are. In either case, there is a more relevant problem. The messenger-shooting problem is also problematic for imperative views.

Imperativists cannot account for the intrinsic badness of an unpleasant pain either. They are also committed to saying that taking painkillers is non-rational. Just receiving an order is not good or bad in itself. It might be annoying to receive an order, but this not so merely by virtue of it being received. It might bother us to receive an order because receiving an order might cause something else that bothers us; it is instrumentally annoying. A command might interrupt us from doing something else, but it is the interruption that is bad, not the order in itself. If the messenger came to the king and commanded him to do something to stop people from dying in the kingdom, it would still be undesirable for the king to shoot the messenger as a way of dealing with the command. The mere fact that the command had been given wouldn't be a normative reason, a good reason, for terminating it or the messenger. The king might have a good reason to send a doctor to deal with the content of the command, especially if the messenger has some form of credibility or authority. However, simply receiving a command is not bad in itself. If everything remains the same, i.e., all else being equal, we are not worse off just by receiving a command or better off by not having a command. Taking a painkiller would be equivalent to shooting the messenger. Taking painkillers is rational; we have good reason to take them to stop feeling an unpleasant pain. However, for imperativists it would also be non-rational to perform such experience-directed actions.

1. 4 Solutions to the messenger-shooting problem

Both imperativists and representationalists have offered possible solutions to the messenger-shooting problem, as I will show in this section. These solutions have taken two main forms: i) to argue that it is rational to take a painkiller due to *instrumental reasons* or ii) that it is rational to take painkillers in virtue of *another mental state* that is not the pain. I think that that both strategies are unsatisfying. The former option will be unable, *in principle*, to capture the non-instrumental badness of pain; the latter option misses the main point of a content account of pain, since it does not explain unpleasant

pain's badness in terms of mental content. I will discuss each of these answers to the messenger-shooting problem.

1.4.1 Pain as instrumentally bad

Let me start with the former strategy. As I said, there is a very strong intuition confirmed by our lived experience and by our behaviour in relation to unpleasant pains, that is, having an unpleasant pain is bad in itself. More specifically, an unpleasant pain is bad because it is unpleasant. The first strategy to deal with the messenger-shooting problem mainly consists in denying that pain is bad in itself. According to this solution, having an unpleasant pain is only instrumentally bad. The reason in virtue of which one takes a painkiller, for example, is not because it is bad in itself to have an unpleasant pain experience. It is something else that justifies this type of bodily-directed and experience-directed actions. What could this instrumental reason be? Imperativists have proposed that what is ultimately bad about an unpleasant pain is that it is interfering with other activities, and it is because of this interference that it is rational to perform experience directed behaviour such as taking a painkiller (Klein, 2015b; Klein & Martínez, forthcoming; Martínez, 2015).

According to imperativists, we can account for taking painkillers even without unpleasant pain being bad in itself. Instead, it is an instrumental reason that explains why we take painkillers. We take a painkiller in virtue of the fact that an unpleasant pain is interfering with our mental to-do-list of other activities. In other words, the pain command is interfering with other commands and this interference is what justifies our experience-directed actions. It is desirable to take a painkiller because the mental interference is bad. If this is correct, this means that the only good reason to take painkillers is to prevent the unpleasant pain command from interfering with your other activities. Once more, if we accept this, there is a good reason to take a painkiller to shoot the messenger, but the reason is not that the message is bad in itself. If we use again the kingdom analogy, the idea would be that the king may shoot the messenger, but it is not undesirable to act in this way. The reason behind killing the messenger is that the message is interfering with other activities that the king has to do. It is in virtue of such interference that the king has a reason and can even be justified to shoot the messenger. I do not think, however, that this solution from imperativists is satisfactory.

Let us consider and contrast two different scenarios. In the first scenario you have to do many things throughout the day: prepare coffee in the morning, take your children to school, go to work, prepare dinner at night, etc. In the second scenario you have to do the exact same things, but you also have a persistent toothache; you are still able to perform all the activities that you were supposed to do and no one could notice the difference in your performance. Which of these two situations is worse? The latter one is worse, I take it, and you would be better off if you could take a painkiller. If this is correct, it shows that you have a good reason to take a painkiller that is not merely in virtue of the pain interfering with your activities. These two scenarios seem clearly possible to me; the fact that the latter scenario is worse shows it is rational to take a painkiller for non-instrumental reason. I think that it is true that we might have good instrumental reasons to take a painkiller - not having an unpleasant pain could be desirable for many reasons - but that is not what is at stake. Even if it is true that an unpleasant pain is bad because it interferes with other activities, this does not show that an unpleasant pain is not also bad in itself. This example shows that we have a reason to get rid of pain for its own sake and, as I argued in section 1.3 of this chapter, content theories are not able to capture this feature.

Let us consider another case. Suppose that you are being tortured with unpleasant pain. There's no way to escape. This, I think, is a very bad situation to be in. Namely, you would clearly be better off without suffering this unpleasant pain, all else being equal. If you could take a pill that stopped you from having unpleasant pains, I think that you would have a good reason to take it. According to imperativists, the reason behind you taking that pill is that in this way such unpleasant pain produced by being tortured would not be interfering with you mental to-do-list. I am not denying that this might be true, there are many other things that you would rather do instead of being tortured. However, what I am trying to point out is that it seems silly to say that the *only* reason why it is bad to be tortured is because the unpleasant pain that you are experiencing is interfering with other things that you would like to do. Having an unpleasant pain is instrumentally bad, but that does not undermine that it is also bad in itself.[11]

I think that the problem for imperativists in accounting for unpleasant pains' badness is structural. The imperativist has to show that an unpleasant pain is not bad in itself, i.e., that the apparent non-instrumental badness consists really in some instrumental reason. The previous two scenarios are meant to show that pain is indeed bad in itself and not only

[11] Bain (2017) presents a similar intuition.

instrumentally bad. That is, having a toothache is bad even when it is not interfering with our activities. I do not think that we can deny that an unpleasant pain is bad in itself. Any attempt to explain non-instrumental badness by appealing to instrumental reasons is doomed to failure. If we accept that something is non-instrumentally bad, we will not be able to account for such badness in virtue of something instrumental. That is, if something is non-instrumentally bad, it will remain bad regardless of any instrumental badness that it might entail. I think that the examples discussed above encourage us to think that unpleasantness is non-instrumentally bad.

The imperativist could then say that we are simply begging the question. Given that we assumed that an unpleasant pain is non-instrumentally bad, then the instrumental explanations won't convince us. The imperativist could argue that, precisely, we take a painkiller only in virtue of instrumental reasons; thinking otherwise is *petitio principii*. But we are not begging the question, I think. Even if an unpleasant pain is not interfering with your mental to-do-list, it would still be bad to be in that state. Unpleasant pain's felt badness does not rely on you being able or not to perform or plan other activities, although it is worse if you are having an unpleasant pain and the pain also interferes with your to-do-list. Even if we followed the imperativists and accepted that we only take painkillers in order to avoid mental interference, we will have to face a point in the justification where something is bad in itself. The imperativists then would have to explain, for example, why it is bad to have a mental state such as pain interfering with a mental to-do-list. Is mental interference bad in itself or is it bad because of something else?

In either case, I think that the felt unpleasantness of pain seems like a very promising candidate for something that is bad in itself, and as long as this is the case content theories won't be able to explain this feature of an unpleasant pain by appealing to the content that is meant to account for the experience. The examples discussed are meant to show that we must accept that an unpleasant pain is non-instrumentally bad. All in all, if we are justified to take a painkiller in virtue of the felt badness of unpleasant pain experiences, this shows that the experience of an unpleasant pain is non-instrumentally bad. Any attempt that tries to explain this felt non-instrumental badness via instrumental reasons will fail.

1.4.2 Pain as bad in virtue of other mental states

Representationalists have offered two main strategies in order to provide an answer to the messenger-shooting problem by appealing to other mental states. It is rational to take a painkiller in virtue of i) a negative emotion directed at the unpleasant pain (Boswell, 2016) or ii) because we have a general aversion to the unpleasant pain (Cutter & Tye, 2014). In both cases we justify taking a painkiller by referring to another mental state that is not the unpleasant pain itself. Boswell's solution is problematic in a similar way to the imperativists arguments. That is, it fails because he is trying to account for experience directed behaviour by referring to instrumental reason. The second option is better, I think, since it does not appeal to instrumental reason. However, I also think that it is problematic since it can only explain the motivation to take a painkiller but it cannot account for the justification to do so.

Let us start with the first proposal. Bowell's solution is based on empirical models of pain according to which a normal pain experience is constituted of three elements: i) a raw pain sensation (that is not good or bad), ii) an affective component (where pain gets its usual felt badness), and iii) a negative emotional component (Fields, 1999; Gracely, 1992; Price, 2000). Following this model, it is in virtue of the emotional component that it is rational to take a painkiller. The reason why we take a painkiller is to stop having the negative emotion directed at the unpleasant pain. If you experience anguish or anxiety due to the felt badness of your pain, then you have a reason to take a painkiller; it is desirable for you to stop the pain because this prevents you from experiencing anguish or anxiety. We are supposed to explain that one takes a painkiller because it is a way of preventing having a negative emotion.

This solution is not going to work, however. It has the same problem that I previously discussed in reference to explaining unpleasant pain's felt non-instrumental badness by appealing to instrumental reasons. As I discussed earlier, these types of non-instrumental explanations are doomed to failure if we think that pain's felt badness is also non-instrumental. That is, if we think that pain's felt badness is bad in itself, then it cannot be that we are only justified to take a painkiller because of the bad emotional response that pain causes. In order for this solution to work, we would have to show that an unpleasant pain is actually not bad in itself. Could it be that taking into account this empirical model shows that we only have instrumental reasons to take painkillers? I don't think so. In short, you could have an unpleasant pain with no negative emotional response and that would still be bad.

Again, it seems that there is quite a strong intuition that experiencing an unpleasant pain is not bad merely because of the emotional response that it typically provokes. Having a negative emotion directed at your unpleasant pain is *worse* than only having an unpleasant pain, but this does not show that experiencing an unpleasant pain is bad only in virtue of the negative emotional response that pain elicits. That is to say, Boswell can account for some other reasons why we might take a painkiller. It can be because an unpleasant pain interferes with other activities, as imperativists propose, or it could also be because this prevents us from having other negative emotions. That said, this does not show that these are the only reasons why we take a painkiller, nor does it show that these are the main reasons. Showing that there are instrumental reasons to take a painkiller does not entail that there aren't also non-instrumental reasons. Moreover, Boswell's proposal does not show that a representational account can explain the non-instrumental badness of pain by appealing to the content that constitutes the unpleasant experience.

Let us now have a look at Cutter and Tye's strategy. First, this solution avoids the shared problem of the previous possible solutions. Cutter and Tye explain that we have a reason to take painkillers because we have *an aversion* to unpleasant pains. It is in virtue of this aversion that we can explain why we take a painkiller when we have an unpleasant pain. This aversion should not be merely explained as an instrumental reason, of course, if we want to avoid the type of problem that we have encountered previously. This aversion strategy escapes such a problem. This is because it is not the badness of the aversion itself that instrumentally explains the badness of the unpleasant pain. Instead, having such an aversion to an unpleasant pain experience explains that the unpleasant pain is bad. What do they mean by 'aversion' then?

A good way of making sense of 'aversion' is to put it in terms of a general con-attitude. The aversion is a desire for the unpleasant pain not to occur. In other words, unpleasant pain's non-instrumental badness can be explained by appealing to the desire to not have an unpleasant pain. This is not a form of instrumentalism because we are not explaining the badness of an unpleasant pain in terms of the badness of the desire. It is neither a desire to *not have an unpleasant pain because in this way we will avoid something else that the unpleasant pain may cause*. It is a desire directed at the unpleasant pain for its own sake. This aversion can be understood as a non-instrumental desire to not have an unpleasant pain. When we take a painkiller, the idea goes, this is not as a way of preventing having this desire; instead, this desire is constitutive of the badness of experiencing an unpleasant pain and it is in virtue of this desire that we can explain, for instance, our motivation to get

rid of an unpleasant pain experience. The desire is what accounts for an unpleasant pain being bad and we can render actions intelligible by appealing to this desire.

Cutter and Tye's view about what constitutes the unpleasantness of pain is a form of evaluativism, such as BAP. They maintain that the content of an unpleasant pain and particularly of its unpleasantness is evaluative. This is not crucial to their solution to the messenger-shooting problem, however. The important thing is that we account for the reasons to take a painkiller by appealing to a non-instrumental mental state distinct from the unpleasant pain. When it comes to experience-directed behaviour such as taking painkillers, one may argue that these actions are not explained in virtue of the unpleasant pain's content, but in virtue of other mental states directed at the unpleasant pain experience. Notice that this is not an instrumental explanation: according to this solution to the messenger-shooting problem we do not take painkillers in order to avoid such desires, the fact that we have these desires is what is meant to account for our motivating and justifying reasons to take painkillers. These non-instrumental desires are meant to render our actions intelligible and, I take it, to also justify such actions. This, I think, is the best available solution that content theories have proposed by appealing to other mental states in order to deal with the messenger-shooting problem.

There are, nonetheless, at least two problems with this solution. The first one is that this explanation can account for our motivation to take a painkiller, but it is not as clear how it should account for our justification to do so. That is, when we appeal to another mental state to account for why we take a painkiller, we are only making such action intelligible, we are offering a motivating reason, but this does not imply that we are justified to take painkillers, that we have a good reason to do so. If we accept that we have a good reason to take painkillers, content theories have not explained why this is. They have not explained this in terms of mental content and they have not explained this in terms of having a desire to not have an unpleasant pain. Having an aversion to an unpleasant pain does not entail that we have a good reason to take painkillers, that it is desirable. Aversion in itself does not imply justification. For example, we can have an aversion towards many things and the mere aversion does not imply that we have a good reason to act in relation to our aversion. In fact, we can have an aversion towards things that are good for us. For example, we might be averse to going to a chemotherapy session, but, all things considered, it might be good since it is the only way to save our lives once we have cancer. There are a few solutions that a content approach could implement in order to account for normative reason in terms of desire. I think that a straightforward answer is to

appeal to desire frustration. However, I will leave this to the next chapter, where I explain the theories that account for the unpleasantness of pain in terms of desire.

The second issue with appealing to aversion as a way of dealing with the messenger-shooting problem is that the theory is no longer capable of explaining that the badness of an unpleasant pain resides in its phenomenology. Moreover, we are not able to explain the badness of an unpleasant pain purely based on the notion of mental content. This is not particularly problematic in itself; one does not have to explain everything there is about the unpleasantness of pain in terms of content, unless one was trying to actually do so. One of the initial virtues of content theories was that we could explain the phenomenology of mental experience such as an unpleasant pain, and everything that rests on that, by appealing to the notion of content. However, it seems that we cannot do this when it comes to the badness of the unpleasantness of pain. Content theories seem to have to use the notion of desire in order to explain this feature. They have to give up the idea that the phenomenology of an unpleasant pain is bad in itself. In fact, there is a different way to account for the unpleasantness of pain that relies solely on the notion of desire.

Finally, content theories have not addressed an important issue regarding unpleasant experiences, including unpleasant pains: the heterogeneity of unpleasant experiences. One key difficulty for content theories is dealing with the heterogeneity of unpleasant experience. Namely, there seems to be no unitary feeling that *all and only* unpleasant experiences share and in virtue of which they all qualify as unpleasant. In a nutshell, all unpleasant experiences feel unpleasant, but they all feel quite different. Content theories consider that one single type of content accounts for one single way of being unpleasant among all unpleasant experiences. However, there is a strong intuition that there are many ways of being unpleasant, which is something that content theories cannot capture as they have been so far proposed. This problem will become clearer in the upcoming chapters, and I will consider other accounts that have directly addressed this issue. In contrast, content theories treat unpleasantness as if there were a unitary unpleasant phenomenal character that all pains, and all unpleasant experiences, share.

1.5 Conclusion

Even if content theories offer an initially plausible account for the nature of pain and unpleasantness, they also face several problems. Both representationalist and imperativist accounts face a major difficulty. Namely, neither can explain why we are justified to perform experience-directed actions such as taking painkillers in virtue of mental content that constitutes an unpleasant pain. The best available explanation for content theories to account for experience-directed actions is to appeal to a non-instrumental aversion. A good way of making sense of this aversion is to phrase it in terms of desires. However, other philosophers have precisely tried to account for pain's unpleasantness by appealing to external mental states such as desires, and have built a theory of the unpleasantness of pain that has the notion of desire as the core feature of the account.

In the next chapter I will explain what precisely these alternative theories say in order to account for pain's unpleasantness in terms of desire. I will refer to these as the *desire theories* of the unpleasantness of pain. Desire theories have an important advantage over content theories: they account for the heterogeneity of unpleasant experiences. I will go into the details of this problem in the next chapter. In the upcoming chapter I will expose and analyse the desire account of the unpleasantness of pain.

CHAPTER 2: DESIRE THEORIES

2.0 Introduction

In this second chapter I will explain and discuss a different way of accounting for the unpleasantness of pain. This account relies crucially on *desires* in order to explain what constitutes the unpleasantness of pain and of other unpleasant experiences. The general idea is to explain that a sensory experience, such as a pain, is unpleasant in virtue of that sensation being desired not to occur (Armstrong, 1962; Brady, 2017; Pitcher, 1970). It is because we desire a pain sensation not to occur that it qualifies as unpleasant. In order to explain how the desire theories account for pain's unpleasantness, and critique such approaches, I will proceed as follows.

First, it is important to clarify what desire theories take desires to be, and so I will clarify this. Second, I will explain how these theories use the notion of desire to account for the unpleasantness of pain and of other unpleasant experiences. For this, I will focus on Heathwood's (2006, 2007, 2011) proposal, which I take to be a well-developed and detailed version of desire theories. I will argue that desire theories have important benefits in contrast to the content theories from the previous chapter; desire theories are in a better position to explain the nature of pain's unpleasantness in the face of the heterogeneity problem. That is, according to the heterogeneity problem, there is no single unitary unpleasant feeling common to all and only unpleasant experiences, yet content theories seem to take unpleasantness as a unitary feeling. Desire theories can accept the problem heterogeneity poses, and can explain that even if this is the case, all unpleasant experiences qualify as unpleasant not in virtue of a unitary feeling, but in virtue of the desires that constitute these unpleasant experiences.

Third, once the desire theories' proposal is clarified, I will focus on two problems for the theory. The first one is raised by Bramble (2013); he claims that we can have unpleasant *unconscious* experiences and that desire theories cannot accommodate this. Broadly, he argues that we can have unconscious unpleasant experiences but that we cannot have desires about these experiences. If this is correct, then we cannot account for the unpleasantness of these experiences based on the desire theories' proposal. I will argue against Bramble and show that there are various ways in which desire theories can accommodate these cases.

The second problem, which I take to be a fundamental objection to desire theories, is often put as a Euthyphro dilemma: do we desire not to be in pain because pain is unpleasant, or is pain unpleasant because we desire not to be in pain? Desire theories have to take the second horn of the dilemma. I will argue that desire theories have an answer to the dilemma, but such an answer is still problematic. The second horn of the dilemma is, intuitively, implausible. Roughly, desire theories defenders fail to account for why we desire not to have pain sensations, if, according to their own theory, there is nothing bad about pain sensations in themselves. Desire theories can still provide an answer for this critique, by explaining that we have these desires as the result of a process of natural selection. However, the natural selection argument that desire theories can offer is also problematic. There are two main issues with this response. First, even if it offers an *explanation* for us having these desires, and a good instrumental reason to have these desires in order to survive, this explanation still cannot provide a non-instrumental reason for having such desire. Second, there is a much simpler and more intuitive answer: to take the other horn of the Euthyphro dilemma and so defend a theory that takes the unpleasantness of pain as a phenomenal property that is bad in itself. I will explain and analyse this alternative approach in Chapter Three.

2.1 Pain's unpleasantness in terms of desire

2.1.1 Sid-desires

Desires are a fundamental type of mental state.[12] That is to say, a desire cannot be reduced to something more basic in terms of mental states. Desires are, in this way, one of the building blocks of the mental building, as it were. A good way of understanding what desires are is to contrast them with beliefs since these are another type of fundamental mental states. One important similarity between desires and beliefs is that they are both typically intentional. To put it simply: they are *about* something. You might believe that you have a red apple or a desire to have a red apple. Both mental states are similar insofar as they are about having a red apple, but there is a very important difference. The difference between these two types of mental states is often explained in terms of *direction-of-fit*.

[12] For more on the nature of desire see Schroeder (2017, 2004).

Let us take a look at this distinction in order to understand what desires are. The direction-of-fit distinction is attributed to Anscombe (1957, §32). She asks us to imagine a man who goes to the supermarket with a shopping list that he follows to pick up the items he has on this list. In this scenario, there is also a detective who carefully writes down on another list everything the man picks up from the shelves. As Searle puts Anscombe's point:

> [T]he function of the two lists will be quite different. In the case of the shopper's list, the purpose of the list is, so to speak, to get the world to match the words; the man is supposed to make his action fit the list. In the case of the detective, the purpose of the list is to make the words match the world; the man is supposed to make the list fit the actions of the shopper… I propose to call this difference a difference in *direction of fit*. The detective's list has the *word-to-world* direction of fit (as do statements, descriptions, assertions, and explanations); the shopper's list has a *world-to-word* direction of fit (as do requests, commands, vows, promises). (Searle, 1979, pp. 3–4)

The two lists are about the same items, but there is a very important difference regarding what these lists are meant to capture. Similarly, beliefs and desires have a very different nature: whereas beliefs are more like the detective's list with a *word-to-world* direction of fit, desires on the other hand are more like the shopper's list and have a *world-to-word* direction of fit. In other words, a desire is doing what it is supposed to do when it is *satisfied*, when the world is in such a way that it matches with the content of the desire. This is similar to the direction of fit that imperative content has, as we saw in the previous chapter. In contrast, beliefs and other representational states can be *accurate*, when their content matches with the world. Another way of making sense of this is in terms of what ought to change when there is a mismatch between the content of the mental state and the world. When there is a mismatch between the content of the desire and the world, it is the world that ought to change for the desire to be satisfied. In contrast, when there is a mismatch between the content of the belief and the world, it is the belief that ought to change in order to be accurate.

A desire is a fundamental mental state that has the function of being satisfied. Now that it is clearer what a desire is, I can move on to explaining how a desire can be used to account for the unpleasantness of pain. Desire theories offer a reductive account of unpleasantness. Unpleasant experiences are so in virtue of being constituted by two more basic mental states: i) a hedonically neutral sensory experience, i.e., a sensation that is neither pleasant nor unpleasant, and ii) a particular kind of desire not to have this sensory

experience (Heathwood, 2007). An unpleasant pain, for instance, is composed of: i) an *inherently* hedonically neutral pain sensation (i.e. it might be unpleasant but it isn't in itself) and ii) a desire to not have this pain sensation. Even if Heathwood's theory and examples are mainly focused on the reduction of sensory *pleasure* to desire, Heathwood's view is also meant to capture unpleasant experiences including the nature of unpleasant pains. Let us now explain in more detail the two components that constitute an unpleasant pain according to this view.

Heathwood takes pain to be a sensory experience. However, he unfortunately does not explain what exactly a sensory experience is. That said, he does suggest that smell and taste experiences may count as sensory pleasures. But, for the moment, we only need to focus on pain. If pain counts as a sensory experience, I understand that this means that pain sensations have a certain phenomenology, that there is something it is like for someone to have a pain sensation; pain is a qualitative state in the sense that there is a pain *quale*, and when you have a pain sensation you are aware of it, at least typically and under normal circumstances. Further, it is crucial to underline that, according to Heathwood, sensory experiences are hedonically neutral, they are not pleasant or unpleasant in themselves — this is an important feature of the theory and I will show that it is also highly problematic.

The second element of an unpleasant pain is the particular kind of desire that is directed at the pain sensation. The desire, according to Heathwood, must be i) simultaneous, ii) intrinsic, and iii) *de re* about the hedonically neutral pain sensation. I will refer to this as a *sid-desire*. It is because of such sid-desire for a pain sensation not to occur, that the unpleasant pain is unpleasant. The desire is about the pain sensation not to occur *qua* a qualitative phenomenal experience, that is, the desire is for a pain sensation quale not to occur. More precisely, it is a sid-desire about a state of affairs, that is, a state in which one is experiencing a pain sensation that is neither pleasant nor unpleasant. For example, think about a smell like the one produced by coffee. The idea is that this smell is in itself neutral, it is not pleasant or unpleasant. Similarly, when one experiences pain, there is a sensory component that is in itself neutral, but this neutral feeling is desired not to occur and it is in virtue of such desire that the whole experience is unpleasant. I will now explain each of the features of a sid-desire in order to make this clearer.

The desire that constitutes an unpleasant pain has to be *simultaneous* with the pain sensation that the desire is about; both elements have to occur at the same time. Why is this time constraint important? This constraint helps to deal with various possible

counterexamples for desire theories. For example, suppose you have a desire in the morning to have the taste experience of a strong black coffee, but it is not until the afternoon when you finally get your strong black coffee experience, yet the taste experience is disgusting, it is very unpleasant. Should it not be pleasant since you wanted it? Not according to desire theories, because the relevant desire in virtue of which the coffee experience is pleasant or unpleasant has to be co-occurring. Instead, the desire that we are considering occurred many hours before. What is really happening in this scenario is that in the morning you wanted the taste, but then, when you actually had it, you didn't want it. The constituting desire that explains that the taste is unpleasant is the one that is directed at the taste sensation at the same time that the sensation is taking place.

Something very similar can be said for pain. In the morning you had the desire to go the gym, at that moment when you stepped out of bed you wanted to feel the burning pain of exercising through weightlifting. Then when you actually go to the gym in the afternoon and do weightlifting, the experience of burning pain felt in your biceps is unpleasant. Shouldn't the pain be pleasant according to desire theories since you wanted it? Not really, because the relevant desire that constitutes unpleasantness must be simultaneous with the sensory experience. The desire that is going to account for experiences being pleasant or unpleasant must be simultaneous to the sensory experiences. To put it another way, only simultaneous desires are relevant in order to explain what makes the experience hedonic, i.e., the types of desire that account for sensory experience being pleasant or unpleasant are necessarily simultaneous to the neutral sensory experiences to which they are directed. This does not mean that all simultaneous desires are relevant in accounting for an experience being hedonic, but it does mean that the desires that are constitutive of a hedonic experience must be simultaneous to the sensory experience. In this example, desire theories explain that even if you wanted to experience that burning pain sensation *in the morning*, the concurrent desire *in the afternoon* is actually for that pain sensation not to occur.

The desire must also be *intrinsic*. A typical way of understanding intrinsic desire is to put it in terms of desiring something 'for its own sake', as opposed to desiring that something for something else, that is, to have an extrinsic desire. For example, suppose that your mother is very sick and you desire that she gets better. You have a desire about your mother's health improving. If you desire this for its own sake, its means that you desire her health to improve *tout court*, you do not desire this for the sake of something else that you desire. In contrast, if you desire her health to improve in virtue of something else, whatever that might be, then your desire about your mother's health would not be

intrinsic. If you desire it because in this way she will be able to come to your wedding, because in this way she will be able to go on vacation, or anything else, then your desire about her health is *extrinsic*. However, I think we need to be more detailed about what it means to desire something 'for its own sake'. I propose that when *X* is desired intrinsically this means the following:

> *S* desires *X* intrinsically if and only if *S* desires *X* without having this desire in virtue of a relation that *S* represents *X* having with some *Y* that *S* also desires.

In contrast, desiring *X* extrinsically means that:

> *S* desires *X* extrinsically if and only if *S* desires *X* in virtue of a relation that *S* represents *X* as having with some *Y* that *S* also desires.

This, I think, gives a finer grained explanation of what we mean by desiring something for its own sake. That is, when you desire your mother's health to improve intrinsically, this means that you do have this desire without representing the content of your desire as having any type of relation with something else. In this case, you desire your mother to get better without representing her getting better as being connected with something else. In contrast, when you desire that same thing extrinsically, this means that you desire your mother's health to improve in terms of a relation that you take to exist between her health improving and something else that you desire. This improvement, for instance, can be seen and linked to her ability to go to your wedding: you think that if she gets better she can come to your wedding. According to desire theories, the type of desire that constitutes the pleasantness or unpleasantness of experiences must be intrinsic. For example, when we account for the unpleasantness of pain and one desires a pain sensation not to occur, this is in virtue of the pain sensation in itself and not in virtue of any relation that the pain sensation is taken to have with something else.

Given that the desire has to be intrinsic, this implies that the relevant desire is also non-instrumental. Instrumentality is a type of relation, and desiring something instrumentally means that we desire that something in virtue of the represented causal relation between what we desire and something else. Broadly, *X* is instrumental if there is some *Y* such that *X* causally brings about *Y*. An instrumental desire for *X* is a desire for *X* in virtue of *X* being represented as having a causal relation with some desired *Y*. For example, if you want to stop being thirsty and you represent water as a means to stop being thirsty, your desire for water is an instrumental one. You desire water on the basis that you believe, i.e., that you represent, that water has the causal power to make you stop being thirsty, which

is something that you also want. Intrinsic desires cannot be instrumental because they are not held on the basis of a represented relation and instrumental desires are held in virtue of a represented relation, a causal relation to be more precise. This means that the desire that constitutes unpleasantness has to be intrinsic, and hence, it cannot be an instrumental desire.

Let us illustrate this last idea. Suppose that you desire to experience a pain sensation as a means for forgiveness. You want me, for example, to punch you in the face, and thus cause you pain, as you think that this will renew our friendship since you did something bad to me (you lied to me about something important, you stole something valuable from me, etc.). I agree and I punch you in the face and cause you pain. As you experience the pain, you also desire to have that pain since by feeling this pain you hope to be forgiven. Shouldn't the experience be pleasant since you wanted it simultaneously? Desire theories can explain that even if you did want the pain simultaneously, the desire was not intrinsic; it was an instrumental desire. What really happened is that you had a *simultaneous extrinsic* desire for the pain to occur, you had an instrumental desire for the pain to occur; however, you also had a *simultaneous intrinsic* desire for that pain not to occur, your desire for such pain not to occur was not held in virtue of any relation that you thought the pain had with something else, including possible causal relations. This is meant to show that the kind of desire that constitutes the unpleasantness of the experiences is not only simultaneous, but also intrinsic. Being simultaneous and intrinsic are necessary features of the desire in virtue of which we account for the experiences being hedonic.

The third and last feature that desires must have in order to constitute unpleasantness is that they have to be *de re*. A good way of understanding what this means is to refer to the distinction between types and tokens. Whereas to desire *X de dicto* means to desire it understood *as a type*, desiring *X de re* is to desire it *as a token*. Let me illustrate this. Take the sentence "I desire a red apple". This can be understood *de dicto*, which means that the desire would be satisfied by any object that follows under the type *being a red apple*; the desire is about a type of object, not about an object in particular. Now, if we understood the sentence *de re*, this means that the desire can only be satisfied by a specific red apple, the desire would only be satisfied by one specific token of the type being a red apple.[13] The relevant desire that is supposed to constitute unpleasantness has to be *de re*, because

[13] This is not the only way of understanding what *de re* means. Another way of understanding *de re* is in terms of an indexical and *de dicto* as not being indexical. I will not go into details about the various ways of understanding the distinction *de dicto vs. de re* since I do not think it makes an important difference for understanding desire theories' proposal.

one may desire some sensory experience *de dicto* to occur, and yet the whole experience be unpleasant, or, vice versa, desire the sensory experience not to occur *de dicto*, yet the whole experience be pleasant.

For example, suppose that you desire to have the experience of flagellation. In this scenario you desire this experience intrinsically, which implies this is not in virtue of any represented relation between the experience of flagellation and something else — including instrumental relations such as the flagellation causing something else. I insist that this desire to experience flagellation is not seen as a means for something else, such as redemption of your sins; you just want the experience for its own sake. Then, as you have the experience of flagellation, you continue to desire it to occur. This might be psychologically odd, but I do not see why it would be impossible to have such a desire. Once you finally have the experience, it is unpleasant. Shouldn't it be pleasant since you intrinsically and simultaneously desire it to occur? According to desire theories, what is really happening is that even if you do desire the experience *de dicto* to occur, you also desire it *de re* not to occur. In other words, even if you do desire this flagellation pain to occur as a type, you also desire it not to occur as a token. With all this in mind, we can give a precise formulation of what desire theories propose.

Desire theories proposal

> An unpleasant pain experience is unpleasant if and only if it is constituted of two components: i) a pain sensation that is inherently hedonically neutral, and ii) a simultaneous, intrinsic, *de re* desire of that pain sensation that it not to be occurring.

All in all, according to desire theories, an unpleasant pain experience qualifies as such in virtue of it being constituted by these two elements. These are necessary and sufficient for having an unpleasant pain. What it means to have an unpleasant pain experience is to have a sensory pain that is, in itself, neither pleasant nor unpleasant, and to have a sid-desire that is about that specific pain sensation. It is important to point out that it is the compound experience that has the property of being unpleasant and not its individual constituents. The pain sensation is not unpleasant in itself and neither is the sid-desire. Instead, an unpleasant pain is constituted by such a sensation and a sid-desire. Now that we have a clear grasp of the proposal, we can turn to analysing its virtues and vices. I will start with the former.

2.1.2 The heterogeneity problem

Let us explain one of the main advantages of desire theories over content theories. If we are trying to find the best possible account for the unpleasantness of pain and other experiences, a strong reason to opt for desire theories is that they can provide an answer to the heterogeneity problem. The general idea is that, on the one hand, all unpleasant experiences feel unpleasant; however, on the other hand, there seems to be no unitary feeling that all and only unpleasant experiences share. This idea also applies to pleasant experiences and, in fact, most of the literature on the heterogeneity of experience is about pleasant experiences. The idea is similar, that is, there is no qualitative aspect that all and only pleasant experiences share and by dint of which they are pleasant. Feldman (2004) provides a good example of this intuition:

> Reflection on sensory pleasures quickly reveals an enormous phenomenological heterogeneity. Perhaps this can be expressed more simply: sensory pleasures are all "feelings", but they do not "feel alike". Consider the warm, dry, slightly drowsy feeling of pleasure that you get while sunbathing on a quiet beach. By way of contrast, consider the cool, wet, invigorating feeling of pleasure that you get when drinking some cold, refreshing beer on a hot day…[T]hey do not feel at all alike. After many years of careful research on this question, I have come to the conclusion that they have just about nothing in common phenomenologically. (Feldman, 2004, p. 79)

Feldman's intuition is, I take it, that even if we agree that the warm and cool feelings in each of the situations are pleasant, there seems to be no unitary feeling in virtue of which both qualify as pleasant. This intuition is also held for unpleasant experiences. If you think of many of the unpleasant experiences that you might have, such as feeling pain, feeling dizzy, experiencing itching, hunger, thirst, etc., there is nothing phenomenal, no conscious unitary feeling, in virtue of which all and only these experiences can be grouped as all belonging to the same type of experience. Korsgaard (1996) writes along these lines:

> If the painfulness of pain rested in the character of the sensations…our belief that physical pain has something in common with grief, rage and disappointment would be inexplicable. For that matter, what physical pains have in common with each other would be inexplicable, for the sensations are of many different kinds. What do nausea, migraine, menstrual cramps, pinpricks and pinches have in common that makes us call them all pains? (Korsgaard, 1996, p. 148)

It is worth noting that Korsgaard is using the word 'pain' here in a very loose sense. That is, nausea is not a pain as a headache is, most would agree. Many would agree too, I take it, that feeling nausea is unpleasant, but it is not a pain as the feelings of being cut or burn are. Furthermore, pinpricks and pinches are stimuli that *cause* pain, but they are not pains themselves. The word that she should be using here is, I think, "unpleasant". That is, all these unpleasant experiences are extremely diverse and they do not seem to have one single phenomenal aspect that unifies them all and by dint of which they all qualify as unpleasant. What this suggests is that being unpleasant is not a shared qualitative feature among all and only unpleasant experiences; being unpleasant, this suggests, cannot be explained in virtue of a shared unitary feeling that all and only unpleasant experiences share. There is nothing that feels alike among all and only unpleasant experiences.

The heterogeneity problem

> We have a very strong intuition that i) all and only unpleasant experiences *feel* unpleasant; it is in virtue of feeling unpleasant that they qualify as such. However, after careful introspection, there is also the strong intuition that ii) there is nothing qualitative, nothing phenomenal, *no unitary feeling*, that all and only unpleasant experiences share and in virtue of which they all count as unpleasant.

So how do desire theories confront the heterogeneity problem? I think that the simplest and most elegant way for desire theories to confront it is to say that unpleasantness is not qualitative, it is not something phenomenal: unpleasantness is not a feeling. That is, there is no single unitary feeling that all and only unpleasant experiences have, because unpleasantness is not a feeling. Instead, desire theories can explain that what unifies all unpleasant experiences, or at least all unpleasant sensory experiences, is that they are partially constituted by a sid-desire. What really unifies all and only unpleasant sensory experiences is that they are constituted by the same type of desire, and it is in virtue of this that they all have the property of being unpleasant. More precisely, what unifies all and only unpleasant sensory experiences is that they are partially constituted by a sid-desire for a hedonically neutral sensory experience not to occur. This is not a denial that there is something phenomenal about unpleasant sensory experiences, including unpleasant pains, since these unpleasant experiences are partially constituted by sensory experiences. However, it is a denial that being unpleasant is phenomenal in any way.

So, how can we explain that many unpleasant experiences are felt in different ways, if unpleasantness is not a feeling? Let us focus on the case of unpleasant pains to make this

simpler to clarify: how can different unpleasant pains feel different and still count as unpleasant if unpleasantness is not a feeling? Desire theories' answer for this must be to account for the phenomenological difference by appealing to the differences in how different sensory pains feel. Take for example the difference between the feeling of being cut or burnt. An unpleasant cut feeling and an unpleasant burn feeling are different experiences insofar as they are constituted by distinct sensory components. These diverse sensory aspects, desire theories should explain, are what account for their different phenomenology. The fact that these two experiences are also partially constituted by a sid-desire not to have a pain sensation is what accounts for them being unpleasant. However, the proposal goes, unpleasantness strictly speaking is not something felt. Feelings might be unpleasant, but unpleasantness is not a feeling. The downside of this view is that we have to sacrifice the strong intuition that unpleasantness is indeed something felt, that unpleasantness is a feeling. The upside of the account is that we can explain in a straightforward way what unifies all and only unpleasant pains and unpleasant sensory experiences.

There are several other aspects of unpleasantness that can be illuminated with this understanding of desire theories. Let us start with motivation. If we consider that a mental state is motivational if we can render behaviour intelligible in virtue of such a mental state, then we could explain how unpleasant pains are motivational. For this, we will rely importantly on the fact that unpleasant pains are constituted by a desire, that is, when we are in unpleasant pain this implies that we have a sid-desire because we desire a pain sensation not to occur while the sensation is occurring. For example, some body directed behaviour could be explained in virtue of our sid-desires. That is to say, given that we desire not to have a pain sensation, we might act in order to satisfy this sid-desire by withdrawing from a source of bodily damage, such as moving away from a burning object. By acting in this way, we do something in order to stop having the pain sensation caused by being burned. Our action is aimed at satisfying the content of our sid-desire for that pain sensation not to be occurring. This action is made intelligible in virtue of that sid-desire. We can thus explain what makes unpleasant pains motivational.

Let us now consider unpleasant pain's normative force, that is, the fact that having an unpleasant pain also justifies certain actions; that unpleasant pains provide justification for action, a good reason to act. In order to explain this, it is important to accept that having frustrated desires is bad in itself and satisfying our desires is good in itself *all else being*

equal.[14] That is, you may desire to jump off a high cliff, and it may be a terrible idea to act in order to satisfy this desire *all things considered*, given the bad consequences of this action. However, everything else being equal, it would be good to satisfy even this desire, since desires are mental states that are ultimately meant to be satisfied. Desire satisfaction is good in itself because, all else being equal, it is better to have satisfied desires than frustrated ones. Desire frustration is bad in itself, because all else being equal it is worse to have a frustrated desire. If we accept this, we can then show how desire theories can explain why unpleasant pains also provide us with good reasons to act. That is, given that it is bad in itself to have frustrated desires, and that unpleasant pains are constituted by frustrated sid-desires, then it is bad in itself to have unpleasant pains. We are, therefore, justified to act in order to not have frustrated sid-desires. The bare fact of having an unpleasant pain implies having a frustrated desire, and this in itself provides us with a good reason to act in order to stop having a frustrated desire.

This understanding of desire theories is, I think, quite straightforward, and, even if it goes against the strong intuition that unpleasantness is something felt, this approach offers a simple answer to the heterogeneity problem. That is, we do accept that various unpleasant experiences feel very differently, yet they are all unpleasant. How can we explain this? Desire theories' answer is simple. Unpleasantness is not a feeling. It is important for desire theories to claim that unpleasantness is not a feeling, because otherwise they would not have given an answer the heterogeneity problem. If desire theories took unpleasantness to be a unitary feeling that is explained in terms of sid-desires, then this feeling does not seem to present in all and only unpleasant experiences. Desire theories have to give up a very strong intuition that unpleasantness is something felt, but they can nicely deal with the heterogeneity of unpleasant experiences. Moreover, even if desire theories do not appeal to any unpleasant feeling, they can still accommodate the motivational force and normativity of unpleasant pains. However, even if this is a promising theory of the unpleasantness of pain, I will now show that it is also problematic.

[14] This is the strategy that Cutter and Tye (2014) may use when they introduce the notion of aversion in order to explain the badness of an unpleasant pain. They could claim that the frustration of such desire provides a normative reason. However, in contrast to their approach, desire theories account for both the unpleasantness of pain and also for the badness of pain in terms of desire. In this sense, desire theories are a more elegant theory than Cutter and Tye's, since they do not rely on mental content in order to explain unpleasantness.

2.2 Problems for desire theories

2.2.1 Unconscious unpleasant experiences

Bramble (2013) raises a problem with desire theories. The argument is that i) we can have unconscious unpleasant experiences, yet ii) we *cannot* have desires about unconscious experiences and, therefore, iii) what accounts for the unpleasantness of these unpleasant unconscious experiences cannot be a desire. If this is correct, then desire theories are mistaken. Desire theories are supposed to give an explanation of unpleasant experiences, yet they cannot explain why unconscious experiences are unpleasant. I first describe a case that Bramble uses to illustrate an unpleasant unconscious experience, and explain why, according to him, desire theories cannot account for the unpleasantness of this experience. I think that there is a clear strategy to deal with Bramble's critique: to show that we can have desires about unconscious unpleasant experiences. That is, even if it is true that unconscious experiences can be unpleasant, we can have desires about these experiences and, therefore, desire theories can offer an explanation for these cases too.

Haybron (2008) offers an example of an unpleasant unconscious auditory sensory experience. Given that desire theories are mostly focused on sensory experiences, this seems like a good possible counterexample. Bramble uses Haybron's case as an example of an unconscious unpleasant experience, i.e., an unpleasant experience that we might be unaware of. The example is the following:

> Perhaps you have lived with a refrigerator that often whined due to a bad bearing. If so, you might have found that, with time, you entirely ceased to notice the racket. But occasionally, when the compressor stopped, you did notice the sudden, glorious silence… In short, you'd been having an unpleasant experience without knowing it. (Haybron, 2008, p. 222)

According to Bramble, this example shows that desire theories are wrong. "This is because *one can hardly have the relevant kind of attitude (be it disliking, not wanting, disvaluing, or whatever) toward an experience that one is entirely unaware of*." (Bramble, 2013, p. 204) If we are *entirely unaware* of an experience, I take this to be an unconscious experience. If we are unaware of the auditory experience produced by the refrigerator, yet that experience is unpleasant, and according to Bramble we cannot have an attitude such as a desire about that auditory experience because we are unaware of the experience, as in, it is an unconscious experience, then what accounts for this experience being unpleasant cannot be a desire. Desire theories are supposed to capture the unpleasantness of

experiences, especially of sensory experiences like an auditory one, yet desire theories are unable to explain what makes this unconscious auditory experience unpleasant. Since we cannot have the relevant kind of attitude towards this auditory experience, i.e., a desire, then it cannot be that a desire constitutes its unpleasantness.

Even if we accept that we can have unpleasant unconscious experiences, I think that Bramble's argument is not convincing. I do not see why we couldn't have desires about such unpleasant experiences.[15] These desires, in fact, could be had both at the conscious and unconscious level, i.e., we might be aware of having these desires or not. What does Bramble mean by having an experience that we are not aware of? He could say that one is unaware of an experience, i.e., that an experience is unconscious, when one has such experience without noticing it, without directing any particular attention towards the experience. However, one *could* make the experience conscious, if one directed one's attention to it. What does it mean to bring one's attention to an experience? I think that Bramble could argue that one way to account for this is in terms of having a meta-mental state, i.e., a mental state about another mental state. That is to say, the moment that we have another mental state about the sensory experience, this means that the sensory experience becomes conscious. If this is correct, then Bramble could argue that we cannot have desires about unconscious sensory experiences, since the very moment that we have the meta-mental state then the experiences stops being unconscious.

For example, when you are driving a car, you might see and hear many things while thinking about something else —you are driving back home after receiving really bad news, say, and most of your attention is directed at the recent news. After reflection, you cannot remember what you saw or heard while driving, you were not aware of these visual and auditory experiences even if you had them. Similarly, you might try to stop smoking, eating junk food, or drinking alcohol, but if there is something to smoke, eat, or drink in front of you, you might find yourself smoking, eating, or drinking without noticing it. You started smoking, eating, or drinking without being aware of it. I think this sense of being unaware of an experience is plausible for Haybron's fridge scenario, in fact. Bramble could argue that, in this sense, one might have an unpleasant auditory experience without being aware of it, because one has the experience without directing one's attention to it. But then, at the very the moment that you direct your attention to the experience, you become aware of the experience.

---

[15] Heathwood (forthcoming) has recently offered a similar answer. He argues that the type of example that Bramble has in mind involves a *weak sense* of being unaware of a sensation, and, in this weak sense, we can have desires about experiences we are unaware of.

This does seem plausible and it is line with a common understanding of consciousness.[16] Bramble could say something similar about sensory experiences and desire. The moment you have a desire directed at another mental state, such as a sensory experience, the sensory experience becomes conscious and, therefore, we cannot have desires about unconscious unpleasant sensory experiences. Once you have another mental state about your smoking, eating, or drinking experience, then you become aware of such previously unconscious smoking, eating, or drinking experiences. Similarly, the moment you desire not to be having these smoking, eating, or drinking experiences, you become aware of the fact that you are having such sensory experiences, and these smoking, eating, or drinking experiences become conscious. This would show that Bramble is right, that is, we cannot account for the unpleasantness of unconscious sensory experiences by appealing to the desire theories' approach.

However, I don't think that Bramble can use this strategy. The reason is that having desires about sensory mental states does not guarantee that the latter are conscious. In other words, we can indeed have desires about unconscious sensory experiences. There are various ways in which this might be possible. First, it could be that both the desire and the sensory experience are unconscious. If Bramble accepts the existence of unconscious sensory experiences, it should not be strange to accept that there are unconscious desires. If this is correct, then the whole process of having a sensory experience and desiring for the sensory experience not to be occurring could be unconscious. When you are driving while having unconscious sensory experiences, you could also have unconscious desires about them. When you are having unconscious smoking, eating, or drinking experiences you could also be unconsciously desiring not to be having these experiences. When you have an unpleasant and unconscious auditory experience produced by the refrigerator, desire theories could explain that this counts as unpleasant in virtue of your unconscious sid-desire for this experience not to be occurring. I do not see why this solution isn't available for desire theories. We can have unconscious desires about unconscious sensory experiences, without being aware of any of these mental states.

What I believe that Bramble is trying to point out is that a desire cannot come into existence, into being, that we cannot form or produce a desire, if it is about something we are completely unaware of. He might think something similar about beliefs: how could a belief come into being if it is about something that we are not aware of as existing *at all*?

---

[16] See for example the higher-order-theories of the mind. In this approach, broadly speaking, a mental state is conscious only when one is aware of such a mental state; one might only be aware if one has a mental state M_1 if there is another mental state M_2 directed at M_1 (Rosenthal, 1986).

This might seem paradoxical because forming a belief about *X* seems to imply that we are somehow aware of *X*. I think it might be rare to form beliefs about something that we are not aware of, but it is neither impossible nor paradoxical. Why couldn't beliefs or desires come into being if they are about things that we are not aware of? We could, I think, produce beliefs or desires about things we are not aware of if the whole process takes place unconsciously, as the previous examples are meant to show. That is, the whole process of belief or desire production can occur unconsciously. If this is right, then Bramble has to accept that we can have desires about unconscious sensory experiences.

Let me motivate the idea of forming and holding desires unconsciously. If we bear in mind that desires can be dispositional, i.e., that we still have certain desires, such as standing-desires, even if we are asleep or in a coma, this means that our desires can be satisfied while we are unconscious. For example, if you have the desire to win the lottery and you actually win it while you are sleeping, your desire is being satisfied as you sleep. That is, if we can hold a desire unconsciously, without being aware that we have such desire, this suggests that various processes concerning desire can take place unconsciously. If this is the case, then desires could also be unconsciously produced. Going back to Haybron's case, there might be two possible scenarios. First, you have a conscious auditory experience together with a conscious desire and, at some point, you become unaware of having both of them: you now hold an unconscious desire about an unconscious experience. Second, you are aware of the auditory experience, then you become unaware of this experience and, unconsciously, you produce a desire for that auditory experience not to occur: you unconsciously form a desire about an unconscious experience. I see no reason why these could not be available explanations for desire theories. There is nothing about the notion of desire production that requires that the process occurs consciously and even if there was, we could still have unconscious desires about unconscious sensory experiences. Desire theories can account for Haybron's case.

If this is not enough to convince you of desire theories' capacity to deal with unconscious unpleasant experiences, there is one last possibility: one could have conscious desires about unconscious sensory experiences. You are aware of having a desire but you cannot access what it is about. I believe something similar can happen with other mental states. For instance, the tip-of-the-tongue phenomenon could be explained similarly. When you know the name of someone, or some place, but you find yourself incapable of accessing that name, you know that you know, but you can't access what you know, i.e., you cannot manage to become aware of the content of your own belief even if you try to direct your attention at it. If we accept this understanding of tip-of-the-tongue beliefs, I think we

could also say something similar for *tip-of-the-tongue desires*. That is, we are aware of having a desire, but without knowing what it is that we desire, what our desire is about. You have a certain feeling of frustration, you know there is something missing, but you don't know what it is, that is, you have a tip-of-the-tongue desire. You might know that there is something that you want, but without knowing what it is. This again is meant to show that we can have desires about unconscious mental states.

How would a tip-of-the-tongue desire apply to Haybron's example? You hear a sound, you have an auditory experience, and you also have a sid-desire for that sound experience not to occur. You then focus your attention on other things: cooking, eating, talking with your family, thinking about things you have to do, etc. As a consequence of these other activities, you stop being aware of the auditory experience. All this time you still have the desire for the auditory experience not to be occurring. At some point, you know there's something annoying happening, you know that you want something but don't know what it is that you want. Then, suddenly, when the refrigerator's compressor stops, you realise what was annoying you. Once the auditory experience stops you realise that that is what you wanted, you wanted that glorious silence: your frustrated desire was the desire for the auditory experience not to be occurring.

In conclusion, Bramble's argument against desire theories is not sound. Even if we accept that we can have unconscious unpleasant experiences, i.e., that we can have an unpleasant experience without being aware of it, I think that desire theories can account for this. This is because, contrary to what Bramble argues, we can have desires about unconscious sensory experience in many different ways. We can have desires about unconscious sensory experience, consciously or unconsciously.

2.2.2 The Euthyphro Dilemma

The second problem for desire theories takes the form of a version of the Euthyphro dilemma: is pain unpleasant because we desire not to have it, or do we desire not to have pain because it is unpleasant? Broadly, whereas desire theories endorse the former horn of the dilemma, the latter seems to be much more intuitive yet unavailable for desire theories. This is why this is a dilemma for desire theories; that is, there are two options, but one is problematic and the other unavailable. It is more intuitive that we desire not to have a pain in virtue of pain being unpleasant, than thinking that pain qualifies as unpleasant merely in

virtue of our desire for the pain not to occur. Desire theories can offer an initial solution to the dilemma. I will show, however, that this first solution will turn out to be problematic.

> Do we desire pleasant sensations because they are pleasant, or are pleasant sensations pleasant because we desire them? The reductor of sensory pleasure to desire responds, "Yes and Yes." Yes, we desire pleasant sensations because they are pleasant (in other words, we desire them in advance because we know we will be desiring them when we get them). And Yes, pleasant sensation [*sic*] qualify as pleasant in virtue of the fact that they are intrinsically desired. (Heathwood, 2007, p. 39)

Heathwood acknowledges the problem for his account and offers a solution to the dilemma. His solution to the Euthyphro dilemma is meant to apply to pleasant sensory experiences, but his 'Yes and Yes' solution can easily be adapted for unpleasant pain. That is to say, yes, we desire not to have pain in advance because we know that we will desire not to have it when we get it, and yes, pain qualifies as unpleasant in virtue of the sid-desires that we get when we have the pain sensation. In order to understand Heathwood's answer, we first need to point out that he takes sensory experiences to be the bearers of the property of being pleasant and, similarly, in his view it is the pain sensation that is unpleasant. This is different from what I take to be the desire theories' proposal. The way I understand desire theories' view will in fact be more helpful to deal with the Euthyphro dilemma. There is an important similarity and difference between the desire theories' view and Heathwood's. The similarity is that in both views we explain the unpleasantness of pain in virtue of sid-desires, i.e., simultaneous, intrinsic, and *de re* desires for the hedonically neutral pain sensation for it not to be occurring.

However, the difference is that whereas Heathwood takes the pain sensation to be the bearer of the property of being unpleasant, I think that it is better to take the compound, constituted by the sid-desire and the pain sensation, as the bearer of the property of being unpleasant. This is what I take to be desire theories' proposal. This nuance is borrowed from Brady (2017). He proposes that it is better to understand the whole compound as the bearer of unpleasantness, instead of the pain being the bearer of the property. I will explain the benefits of this view after discussing how it is problematic to understand that the pain sensation is the bearer of the property of being unpleasant *à la Heathwood*.

Let us make sense of Heathwood's response to the Euthyphro dilemma. Heathwood argues that, yes, we desire not to have an unpleasant pain sensation in virtue of its unpleasantness. Heathwood is providing a reason for having a desire not to have

unpleasant pain experiences. More precisely, the reason why we desire not to have an unpleasant pain is because they are unpleasant. This is a justification for the desire to not have an unpleasant pain. The pain being unpleasant is a normative reason to desire not to have such pain. It is desirable to have such desire about an unpleasant pain because of the unpleasantness of the pain. That is, Heathwood is showing that we have a good reason to desire not to have unpleasant pains, in the same way as the first horn of the dilemma points out, i.e., we desire not to have unpleasant pains because they are unpleasant. However, it is not this desire - the desire regarding the unpleasant pain - that the Euthyphro dilemma is aimed at.

I think that we can better understand Heathwood's answer to the Euthyphro dilemma in terms of two different desires: a desire D_1 *in virtue of which* the unpleasant pain sensation is unpleasant, and a desire D_2 that is directed at that same pain sensation *because* the sensation is unpleasant. Heathwood can provide a justification of our D_2 desires; we can explain why it is desirable to have these D_2 desires. We have a good reason to have a D_2 desire because this desire will stop us from having an unpleasant pain. It is bad to have an unpleasant pain, we can explain, because it implies having a frustrated D_1 desire. Since desire frustration is bad in itself, this provides a good reason to have a D_2 desire. Desire frustration is intrinsically bad, since the nature and function of desire is to be satisfied. Put in another way, all else being equal, we are always better off with a desire being satisfied rather than unsatisfied, and always worse off with a desire being unsatisfied rather than satisfied. If we accept this, then we are justified in D_2 desiring not to have an unpleasant pain sensation, because in this way we won't have a frustrated D_1 desire.

However, Heathwood still has a problem. He has not provided any account for the D_1 desire regarding the hedonically neutral pain sensation. That is, when we ask through the dilemma if the pain sensation is unpleasant in virtue of our desire not to have it, we are still missing an account for this D_1 desire that is meant to account for the unpleasantness of the experience. In short, Heathwood has not explained why one would take the second horn of the dilemma. He says that, yes, unpleasant sensations qualify as unpleasant in virtue of the fact that they are intrinsically desired not to occur, but we have not given yet any normative or motivating reason not to desire a pain sensation that is not unpleasant in itself. Heathwood has not provided a reason for the D_1 desire. That is to say, so far there seems to be no reason why someone would desire not to have a hedonically neutral pain sensation.

The Euthyphro dilemma remains problematic. We can rephrase it more precisely: do we have a D_2 desire not to have a pain sensation because the sensation is unpleasant in itself, or is the pain sensation unpleasant in virtue of us having a D_1 desire for the sensation not to be occurring? The first horn of the dilemma seems quite sensible; there is a good reason, a justification, for having this D_2 desire. Heathwood could explain that this D_2 desire is justified by appealing to the avoidance of D_1 desire frustration. Heathwood wants to answer "yes" to both horns of the dilemma, but he has not really been able to say "yes" to both of them. He would say that, indeed, a pain sensation is unpleasant in virtue of a D_1 desire for the sensation not to be occurring, but there is no reason that accounts for one having a D_1 desire. Why would you have D_1 desires about completely hedonically neutral pain sensations? As an analogy, if an experience such as seeing something red is not unpleasant in itself, there seems to be no particular reason for wanting an experience of seeing something red not to be occurring. Heathwood's "Yes and Yes" answer is unsatisfactory because, yes, he can account for *prospective* D_2 anti-unpleasant pain desires, but, no, he cannot account for the *simultaneous* D_1 anti-hedonically neutral pain desires. He has not explained the grounds upon which one would have a D_1 desire about something that has nothing experientially bad for oneself.

There is another strategy available to try to deal with the Euthyphro dilemma. I think this one is successful, as it does not consist in giving an answer to the dilemma but in showing that the dilemma does not apply to desire theories. That is, the strategy is to argue that it is not the pain sensation that bears the property of being unpleasant, but rather the *whole compound* of the hedonically neutral pain sensation plus the sid-desire about that pain sensation (Brady, 2017). This solution offers an explanation of what *constitutes* an unpleasant pain, without claiming that something that was not unpleasant in itself becomes unpleasant in virtue of standing in relation to our desires. What I explained as the desire theories' proposal is precisely this. It is a version that accounts for the unpleasantness of pain in terms of sid-desires, as Heathwood proposes, but from a perspective by which we understand the property of being unpleasant as Brady does. That is, we have a D_1 sid-desire directed at a hedonically neutral pain sensation, and this desire together with the pain sensation bear the property of being an unpleasant pain. Additionally, we also have a D_2 desire directed at the compound. In this version there is no dilemma. If we accept this constitutive explanation, there is simply no dilemma to be raised.

Nevertheless, there is still an unsolved problem. Desire theories' proposal has still not been able to account for the D_1 desire that constitutes an unpleasant pain. That is, even if

there is no dilemma for desire theories, the dilemma allows us to notice that there is a weakness in the theory. We can offer a sensible justification for D_2 desires directed at the unpleasant pain experience, but the D_1 desire remains unexplained in the theory. Unpleasantness seems to have a very tight connection with normativity and justification, yet don't we have any reason at all for having the sid-desire that constitutes an unpleasant pain? If we accept this version of desire theories, *à la Brady*, we can justify the proscriptive D_2 desire to not have an unpleasant pain in terms of avoiding D_1 desire frustration, since an unpleasant pain is still constituted by a frustrated D_1 sid-desire. However Brady's solution has the same problem as Heathwood's when it comes to accounting for the simultaneous D_1 sid-desire that constitutes an unpleasant pain. That is, there is no clear reason for having desires about the hedonically neutral pain sensation.

2.2.3 The evolutionary explanation

Desire theories can try to offer a solution for this. This solution is independent if we take desire theories as Brady or as Heathwood understand the unpleasantness of pain, that is, regardless of whether we think that what is unpleasant is the pain sensation or the compound of the sensation plus a desire. However, given that Brady's explanation avoids the Euthyphro dilemma, I take his as preferable, as I previously explained. The solution is, in any case, to offer an *instrumental normative reason* why we have sid-desires for a pain sensation not to be occurring, even if such pain sensations are not unpleasant in themselves. However, it is crucial to notice that this solution does not provide a *motivating reason* or a *non-instrumental normative* reason why an individual would have such sid-desire. This is a teleological and evolutionary explanation, to be more precise. That is, according to this solution, our sid-desires have a purpose and such purpose is to help us to stay healthy and alive, i.e., these sid-desires are evolutionarily advantageous for us to have. I will refer to this as the *evolutionary explanation*.

> **The evolutionary explanation**
>
> It is useful for creatures like us to sid-desire not to have hedonically neutral pain sensations, because in this way we avoid bodily damage and thus are more likely to survive and stay healthy.

In an evolutionary story of our development, we can argue that what explains why we have sid-desires for pain sensations not to be occurring is that our ancestors were more likely to survive by having such desires and they passed this on to us. We are hardwired in

such a way that we have sid-desires not to have hedonically neutral pain sensations. Given that these pain sensations are often the result of or linked to bodily damage, by sid-desiring not to have these hedonically neutral sensations we are better off, because having these sid-desires helps us to avoid and stop bodily damage. Hall (1989) has a similar idea. He thinks that there is nothing intrinsically bad or awful about pain experiences, but creatures like us have developed to inherently dislike them; "evolution has done its work very well and almost every living creature in the animal kingdom finds the sensations accompanying almost every kind of nociception unpleasant. So goes the evolutionary story." (Hall, 1989, p. 648) However, even if the evolutionary explanation provides a sensible explanation of why we have this kind of desire, it is not proving what I think we really need: an explanation of why an individual would form a sid-desire about a hedonically neutral pain sensation. We still lack a complete understanding of why we have such sid-desires, an understanding of the motivating reasons and the non-instrumental reasons that a sentient being has behind such sid-desires.

For example, when we ask someone "why do you want to marry me?" there could be many explanations for this desire. We could provide motivating reasons for desiring to marry someone, reasons that render having such desire intelligible. Our desire to marry someone could be explained by our belief that this person is intelligent, sexy, reliable, etc., i.e., by our belief that this person is somehow good. The belief about that person being good is a motivating reason to have the desire to marry him or her; having this belief explains why we have the desire to marry that person. Moreover, we could provide normative reasons to desire to marry someone, a good reason to have the desire to want to marry such a person. There might be good instrumental reasons for having the desire to marry someone. If by marrying this person we will be happier, believing that this person is good is a good instrumental reason for us to have the desire to marry that person. By having this desire we will be motivated to marry the person, and thus we will be happier. This last type of explanation is the one that the evolutionary explanation provides. However, the evolutionary explanation does not provide other kinds of reasons that we would also expect pain to provide for our desires about pain: motivating reasons and non-instrumental normative reasons.

According to the evolutionary explanation, having sid-desires not to have hedonically neutral pain sensations is part of a process of natural selection. There might be a good reason to have these desires, i.e., having these desires helped us evolve to survive. It is desirable to have these sid-desires in order to guarantee our survival. However, even if the evolutionary explanation provides an instrumental justification of why we might have sid-

desires not to have hedonically neutral pain sensations, I think that this accounting for sid-desires is lacking something: i) we have not yet explained the *motivating* reason for which we have these sid-desires, and ii) we have not given *a non-instrumental reason* for having these sid-desires. We, as individuals, do not have these sid-desires because we think in this way our species will survive, nor because of the way pain sensations feel. At the end, pain sensations in this view are not unpleasant or experientially bad for the individuals. Moreover, there is a strong intuition that desires concerning pain are non-instrumentally justified and motivated, pain is a motivating reason and a good non-instrumental reason for desiring not to have pain. However, the evolutionary explanation behind sid-desires only offers an instrumental justification of such desires. It is good to have these sid-desires *in order to* survive.

Let us try to make this clearer. Something similar could be said about other sensory experiences, e.g., we could explain why it is evolutionarily advantageous for us to be good at perceiving red objects. However, even if we can give an evolutionary explanation of why creatures that were good at perceiving red objects were more likely to survive, this does not mean that we have provided a motivating reason for these creatures to look at red objects, or a non-instrumental justification of it. Even if it is good for us to have perceptual systems to detect red objects with less difficulty than objects of other colours, this is not a motivating reason for us to look at red objects nor does it mean that we have a good non-instrumental reason to do so. When we offer an evolutionary explanation in these cases, we offer an explanation of the phenomenon, and we can even provide an explanation of why it is desirable for us to have evolved in such a way. However, we have not yet given an explanation of why individuals act in the way they do, i.e., we have not provided a motivating reason or a non-instrumental reason for individuals' behaviour.

Similarly, when we explain that it is good for an individual to sid-desire her own pain sensation not to be occurring, we have not explained the motivating reason or the non-instrumental reason for this sid-desire. We can explain why a volcano explodes, why glucose is needed for certain biological processes, why our visual system evolved for being particularly good at perceiving red objects, and why, according to the evolutionary explanation, having certain sid-desires is advantageous for our survival. However, this does not mean that there is a motivating reason or a non-instrumental justification for the volcano to explode, for glucose being needed for certain biological processes, for being good at perceiving red objects, or for having sid-desires not to have hedonically neutral pain sensations. Desire theories are unable to rationalise the constitutive sid-desires of an unpleasant pain in terms of motivating reasons and in terms of non-instrumental reasons.

This is problematic when we consider that there is a strong intuition that this is precisely the type of reason that pain offers in relation to desire.

Furthermore, the teleological evolutionary explanation for our sid-desires can be problematic in multiple other ways. First, it is not clear that we really need sid-desires about pain sensation for survival, or that it is the best possible explanation of how we prevent bodily damage. For instance, *anti-damage* desires might also play an important role in our survival and might even be more efficacious for protecting our bodies under certain circumstances. If we accept the evolutionary explanation as correct it is because, I take it, it is a good explanation of how we survived. Being a good explanation of our survival would make such an explanation at least more appealing, despite its inability to offer a rationalisation for the motivating reason or for the non-instrumental justification of our sid-desires for a pain sensation not to occur.

However, there are other competing explanations of what we needed to survive. For example, anti-damage desires might have also played an important role. That is, we could explain that we are equipped with desires to avoid damage and that these desires might be even more effective than sid-desires directed at pain sensations, since anti-damage desires address the main issue directly. That is, it is probably more efficacious for avoiding bodily injury to desire not to have bodily injury than desiring not to have a hedonically neutral pain sensation that is highly associated with bodily injury. I won't go into the detail of this, but I do want to point out that even if desire theories can provide an explanation for why we have sid-desires for pain sensations not to be occurring by appealing to an evolutionary tale, the evolutionary explanation is not the only possible account of how desires might have played a role in our survival, which diminishes, I think, the appeal of the evolutionary tale.

Second, desire theories' view of unpleasantness might be too demanding in terms of what is actually needed in order to account for how we protect our bodies. This is because one might argue that the actions to protect our bodies via our pain sensations require beliefs about the connection between the damage and the pain sensations, together with more beliefs about how to deal with such bodily damage in order to stop the pain sensation. This kind of cognitive requirement may appear as too demanding for the kind of creatures that we might want to ascribe unpleasant pains to, together with self-preserving bodily behaviour. These might include, for example, very young infants and other mammals to which this kind of belief is not often attributed. That is to say, desire theories' explanation requires a sophisticated apparatus of knowledge of the connection between pain sensation

and bodily damage, yet it seems unlikely that many animals that do have unpleasant pains and act in order to protect their bodies guided by such experiences, have such type of knowledge. If this is correct, desire theories are not actually offering a good explanation of how sid-desire played an evolutionary role.

2.3 Conclusion

Regardless of the difficulties just mentioned, there is a more straightforward reason to reject desire theories. The other horn of the Euthyphro dilemma is still available. Even if the dilemma cannot be applied to desire theories if we take the unpleasantness to be a property of the whole compound experience, the dilemma gives us the possibility to account in a simple way for the relation between unpleasant pains and desires. We can forget about the sid-desire that constitutes the unpleasantness of pain since we can explain such unpleasantness in different terms. We can simply take the first horn that the Euthyphro dilemma offered. Why do we desire unpleasant pains not to be occurring? Because unpleasant pains *feel* unpleasant and feeling unpleasant is bad in itself. Pain feels unpleasant and this is why we desire not to have it. The fact that unpleasant pains feel bad gives us a motivating reason and a non-instrumental justification to desire these pain experiences not to occur. Moreover, this is also consistent with the evolutionary explanation. It is good for us to desire not to have unpleasant pains, because this is also a good way to avoid bodily damage. In the upcoming chapter I will consider theories that take unpleasantness to be a phenomenal property, without thinking that such phenomenology consists in some form of mental content, as content theories proposed in Chapter One. In the next chapter I will start by analysing the distinctive feeling view of unpleasantness. According to this view, the unpleasantness of pain is a phenomenal property that all unpleasant pains have. In the next chapter I will explain and examine this proposal.

CHAPTER 3: THE DISTICTIVE FEELING THEORY

3.0 Introduction

In this third chapter I will explain, analyse, and critique the distinctive feeling theory. In this account there is something it is like to undergo an unpleasant experience, unpleasantness is a single qualitative property that all and only unpleasant experiences have and by dint of which these experiences qualify as unpleasant. This view is meant to explain how such a phenomenal property, the property of being unpleasant, qualifies certain mental states. According to this view, unpleasant experiences are unpleasant in virtue of having such a phenomenal property. This chapter is devoted to trying to understand what this phenomenal property is, and if we can really account for the unpleasantness of pain and other unpleasant experiences based on this view. I will show that we cannot.

There is an important difference between the distinctive feeling theory and the content theories from Chapter One that should be pointed out. In order to explain this difference, we should acknowledge first a similarity: both types of theories take unpleasantness to be a feeling. However, the difference is that only content theories try to account for this feeling. Content theories account for felt unpleasantness by appealing to the mental content that is meant to constitute such unpleasantness. In contrast, the distinctive feeling theory focuses on describing how to understand this phenomenal property, and how it qualifies certain *sensory experiences*, rather than trying to offer a reductive account of what constitutes such phenomenal property.

I will start by describing the distinctive feeling theory. Once the theory is established, I will explain the first problem within this account. The problem is primarily concerned with how the distinctive unpleasant feeling can be a motivating and normative reason for action. One offered solution is to appeal to normative beliefs about such distinctive unpleasant feeling. However, I will show that this solution is problematic. I argue that in order to account for unpleasantness being motivational and normative, we should understand this property as bad in itself.

Even if the distinctive feeling theory can account for motivation and normativity of unpleasantness, I will show that, ultimately, it is unable to provide a satisfactory explanation of the unpleasantness of pain and of unpleasantness in general. The main issue for this theory is the heterogeneity problem. The distinctive feeling theory establishes that

all unpleasant experiences count as unpleasant in virtue of a shared phenomenal property that all and only unpleasant experiences have. All unpleasant experiences share the same ingredient, as it were, which is a unitary feeling of unpleasantness. However, according to the heterogeneity problem, there is no single qualitative aspect, no single unitary phenomenal property, no single 'ingredient', by dint of which all and only unpleasant experiences qualify as unpleasant. If the heterogeneity problem is correct, as I will show, the distinctive feeling theory is wrong.

Finally, I will consider a possible argument in order to save the distinctive feeling theory from the heterogeneity problem. This solution is based on the idea that if we can make comparisons among unpleasant experiences, this must entail that all unpleasant experiences that are compared share the same unpleasant phenomenal property. I will show that this last attempt is unsuccessful in saving the distinctive feeling theory.

3.1 The distinctive feeling theory

Bramble (2013) defends the idea that pleasant and unpleasant sensory experiences are respectively pleasant and unpleasant in virtue of a *distinctive feeling*. A distinctive feeling is a distinctive quality, a phenomenal aspect in virtue of which a sensory experience qualifies as pleasant or unpleasant. More precisely, I think that a clear way to understand what Bramble has in mind is to say that an unpleasant distinctive feeling is a *phenomenal property* of mental states such as sensory experiences. I think we can understand this feeling being *distinctive* in two main ways: i) in terms of a unitary feeling, i.e., it is the same feeling in virtue of which all and only sensory experiences that are unpleasant qualify as such, and ii) in the sense that it is available to introspection, that is, the distinctive feeling is something that is part of our conscious experience and we can direct our attention to it, at least under normal circumstances. If unpleasantness is a distinctive feeling, it means that it is a unitary, single, and introspectable aspect of sensory experiences. For instance, unpleasant auditory, taste, touch, or pain experiences are all unpleasant because they instantiate the phenomenal property of being unpleasant, i.e., because they share a common unpleasant feeling.

Bramble thinks that sensory experiences can exist without being hedonic, i.e., without being pleasant or unpleasant. He accepts the possibility of pain existing as a hedonically neutral sensory experience, in the same way that one might accept that an auditory experience of hearing a musical note, or a gustatory experience of tasting a flavour, does

not have to be pleasant or unpleasant. I think that a good way of making sense of these cases is that one can have mental states that instantiate the property being a sensory pain, i.e., these experiences feel like a pain, but this pain-like experience does not have to have the phenomenal property of being unpleasant. Note that the property of being unpleasant is a property of being a pain, but not vice versa. In other words, the two properties are not qualifying the mental state at the same level. The property of being unpleasant is dependent on the property of being a sensory pain, but not the other way around. I think that it is better for Bramble to understand the property dependence in this way because this can accommodate how we can have solely sensory experiences without them being unpleasant; however, it is much less obvious that we can have mental experiences that are merely unpleasant, but not sensory or phenomenal in another way. I think that a clearer way of putting this idea is that being unpleasant is a *meta-phenomenal property*, i.e., it qualifies other phenomenal properties such as the property of being a sensory pain experience, a touch experience, an auditory experience, etc. I consider this to be a sensible way to understand the distinctive feeling theory.

However, one of the problems for the theory is to explain how exactly the distinctive unpleasant feeling can explain the motivational and normative force of hedonic experiences. That is, when we have an unpleasant pain, it seems that it is because of its unpleasantness that we can render an action intelligible, i.e., that unpleasant pains are *motivational*. It is by appealing to this distinctive feeling, one would think, that we could explain actions in relation to an unpleasant pain. Moreover, unpleasant pains also give us good or bad reasons to act because they are unpleasant, that is, unpleasant pains are also *normative*: it is desirable to act in virtue of the unpleasantness of pain. However, it is not clear how the unpleasant feeling that Bramble proposes is supposed to capture this aspect of unpleasantness. Let me try to be clearer on why it is not obvious that the unpleasant distinctive feeling explains the motivational and normative force of an unpleasant pain.

There are other distinctive phenomenal aspects of experience that are neither motivational nor normative. If we consider other distinctive experiences that might also count as phenomenal, it is hard to explain why these other distinctive phenomenal experiences are not motivational or normative in themselves, yet the distinctive unpleasant feeling is. Think, for example, about the distinctive visual experience of seeing something red. This experience qualifies as a visual experience of seeing red because it instantiates a phenomenal property, the experience has a distinctive *red-feeling-ness*, i.e., the phenomenal quality common to visual experiences of red objects undergone by normal subjects in normal view conditions. Red-feeling-ness is the phenomenal quality associated

with certain visual experiences, in a similar way that we could talk of unpleasantness when we refer to pain experiences. Red-feeling-ness is not a property instantiated by an object such as an apple. Being red, or redness, is a property of an apple. Red-feeling-ness is the property of a mental state, i.e., the property of a visual experience. Having an experience of something being red is, I think, quite distinctive in the sense that we can introspect that we are having a visual experience of something being red. That is to say, there is something it is like for us when we undergo a visual experience of seeing something red, there seems to be a unifying commonality among all and only the experiences of seeing something red, and this common phenomenology is available to introspection, at least under normal circumstances.

The red-feeling experiences and the unpleasant-feeling experiences seem to be distinctive. There is something it is like to undergo experiences with these properties, and we can identify if we are having a red-feeling or an unpleasant-feeling experience under normal circumstances. However, and this is what is problematic for the distinctive feeling theory, an experience of seeing something red does not seem be a motivating reason or a normative reason in itself — not even a reason that might be overridden by other stronger reasons. The distinctive feeling theory should explain the motivation and justification resulting from experiences being unpleasant, i.e., of experiences instantiating a phenomenal property of unpleasantness. However, it is not clear how instantiating a phenomenal property in itself is motivational or normative. For example, visual experiences instantiating phenomenal properties are not in themselves motivational or normative.

How is it that the distinctive experience of seeing something red is not a motivating or normative reason for action, but that a distinctive unpleasant experience is both motivational and normative? Let me rephrase this, how is it that having a distinctive red-feeling experience in itself does not render actions intelligible, but a distinctive unpleasant-feeling experience does? If this is not obvious, it is even less evident when it comes to normativity: how can we explain that a distinctive red-feeling experience does not give us in itself a good reason to act, but that a distinctive unpleasant-feeling experience does? If they are both mental states instantiating phenomenal properties, why does only one of them have these further characteristics? Sobel (2005), for example, raises this doubt for pleasant experiences. ‘‘Given the historical significance of versions of hedonism that claim a phenomenological commonality between pleasures, it is surprisingly obscure what can be said by way of vindicating the reason-giving status of such states.’’ (Sobel, 2005, p. 445)That is, if pleasant experiences provide reasons for

action in virtue of being phenomenal, it is not obvious how being phenomenal accounts for these experiences providing reason for action. Alston (1967) has a similar worry:

> What we are suggesting to be necessarily true is (P) the fact that one gets pleasure out of x is *a reason for doing or seeking* x…The conscious-quality theory can throw no light on this necessity... Why should it be necessarily true that a certain unanalyzable quality of experience is *something to be sought*? (Alston, 1967, p. 346, my emphasis)

Bramble offers a response to this: "in the normal case of attraction to pleasure and aversion to pain, our attraction and aversion is the product of *normative beliefs*… our attraction to pleasure, and aversion to pain, is the result of *our regarding our pleasures as good and our pains as bad*." (Bramble, 2013, p. 215, my emphasis) There is, I think, a clarification needed of Bramble's answer. When Bramble talks about aversion 'to pain', I think that he is referring to aversion to our unpleasant pain, since he accepts that pain can be, at least in principle, hedonically neutral. Even more precisely, the aversion is towards the unpleasantness of the pain. But leaving these clarifications aside, I think that Bramble's central idea is the following: we can explain the motivational aspect, and even the normative force associated with unpleasant experiences, by appealing to *normative beliefs* about these unpleasant experiences.

Bramble thinks that the difference between a distinctive unpleasant-feeling experience and a distinctive red-feeling experience is that we normally have normative beliefs about the former and not about the latter. It is in virtue of this normative belief that we should be able to account for the motivation and normativity of unpleasant pains then. Note, however, that it is not the unpleasant experience in itself that is motivational or normative, it is rather the fact that we have a belief about it that accounts for motivation. It is because the unpleasant-feeling experience is believed to be bad that we can account for motivation in relation to this experience, Bramble argues. This proposal is problematic in various ways. Let us start with the motivational aspect. According to Bramble, having an unpleasant pain is motivational insofar as we have a normative belief about the unpleasantness of the experience being bad. That is, Bramble could argue that we can render actions intelligible by appealing to such belief. It is not strictly speaking the unpleasantness of pain that is motivational; rather the normative belief about the unpleasantness of pain is the motivating reason to take action.

This solution is problematic, however. First, we could argue that evaluative beliefs like this one - beliefs of something as being bad - are in themselves motivationally inert. That

is to say, these kinds of beliefs by themselves are not enough to render actions intelligible, they need other mental states such as desires in order to account for actions. For example, your belief alone about smoking being bad for your health is not enough to count as a motivating reason, i.e., this belief alone cannot render an action intelligible, one might think. When you are trying to stop smoking, someone offers you a cigarette and you say 'no', we cannot explain this action *merely* in virtue of your belief about smoking being bad, we would also need to attribute to you a desire to stop smoking in order to really be able to account for your behaviour. This would show, against Bramble, that the mere normative belief that he uses to account for action is not sufficient to be a motivating reason.

Bramble could argue that the normative belief is in itself a motivating reason, but that there are other stronger competing motivations, which explains why we might not act in accordance to our normative beliefs, even if we have them. However, the mere normative belief might not even be a motivating reason that could be overridden by other stronger motivating reasons. In contrast, a desire to stop smoking may constitute a motivating reason that could be overridden, i.e., all else being equal and in the lack of other competing and stronger motivations, such a desire would motivate you to stop smoking. This is because desires are typically motivating states, i.e., they have world-to-mind direction of fit, as we saw in Chapter Two. This desire counts as motivational, because we could explain certain behaviour by appealing to such desire as a motivating reason for action in the lack of competing motivations. In contrast, evaluative beliefs do not seem to constitute in themselves motivating reasons for actions, not even as motivations that might be overridden. One could argue that, even if you did not have other competing motivations, the belief that smoking is bad does not constitute in itself a motivating reason to act.

Beyond this, there is one more problem for Bramble if we try to account for motivating reasons in terms of normative beliefs: his proposal looks overly intellectual. If we need to have evaluative beliefs in order to explain the motivation associated with unpleasant pains, this may become too demanding in order to apply it to many cases. For instance, if we think that other animals and young children have unpleasant pains and are somehow motivated by these, we must explain that this is so because of their evaluative beliefs about unpleasant pains being bad. This looks too demanding, that is, dogs and babies don't obviously have evaluative beliefs about their unpleasant experiences - especially if we think that evaluative beliefs require the ability of having concepts - yet they are motivated by these experiences. If this is correct, it cannot be that unpleasant pains are

motivational only via evaluative beliefs, as Brambles proposes. Bramble tries to solve this worry by suggesting that other animals, and probably very young infants, are capable of having some sort of very basic 'insight'. This insight should explain how they could represent their unpleasant pains as being bad; a dog and a baby might not have evaluative beliefs, probably because they lack conceptual abilities, but they have *evaluative insights* about the unpleasant experience being bad. However, I think that this is still problematic. I think that once Bramble introduces the notion of evaluative belief in order to account for actions related to unpleasant pains, he inherits similar problems to the ones that content theories and desire theories have.

Let me start with the problems in relation to content theories. I think the clearest way to explain this difficulty is in relation to normative reasons. That is, not only do we think that unpleasant pains motivate action, but also that they justify action: the unpleasantness of unpleasant pains is a good reason to take painkillers. Even if Bramble only explicitly appeals to normative beliefs in order to account for the motivation associated with the unpleasantness of pain, one would expect that same evaluative belief should serve to account for the normativity associated with unpleasantness. The evaluative belief that Bramble appeals to in order to account for the motivation associated with unpleasant pain should also illuminate, I think, the normativity associated with unpleasant pains. This is because a theory of the unpleasantness of pain should explain how the unpleasantness of pain is both motivational and normative, yet Bramble has not given an explanation for unpleasantness being normative.

Having an evaluative belief of an unpleasant pain being bad should not only give us a motivating reason for action, but also a good reason. However, an evaluative belief that something is bad does not entail a good reason to act, not even as a reason that might be overridden by other stronger ones. In the case of the evaluative version of content theories - the account of representational content defended by Bain (BAP) - *a bodily disturbance* was represented as being bad for oneself. However, this evaluative content does not seem to constitute a normative reason in itself, as I argued in Chapter One. According to Bramble, it is *an unpleasant pain* that is believed to be bad. Even if what is believed to be bad is different in each of these theories, in both cases the representation of something as being bad does not account for the normativity associated with unpleasant pains.

Let me illustrate this. The fact that you represent eating vegetables as being bad for you does not entail in itself that you have a good reason not to eat your vegetables. It is not desirable not to eat vegetables solely on the basis of believing that eating vegetables is

bad. An evaluative representation about something being good or bad does not entail justification in itself, i.e., just believing that something is good or bad does not entail in itself a good or bad reason to act in accordance to the belief. One simple way of understanding why evaluative representations are not normative reasons in themselves is that these representations might be false. Bramble's appeal to evaluative belief has precisely this problem. This is because the nature of representations is that the content of the representational states can be inaccurate. In other words, representing something as being bad does not entail that there is something that is really bad, thus representing something as bad does not entail having a good reason to act in relation to that evaluative representation. The distinctive feeling theory cannot explain why unpleasant pains are normative by merely appealing to the normative belief of unpleasant pains being bad. The distinctive feeling theory claims that our aversion to an unpleasant pain is explained in terms of an evaluative belief, but a mere evaluative belief is not necessarily a normative reason that can explain action in relation to an unpleasant pain.

Moreover, the distinctive feeling theory may also lead to a Euthyphro dilemma, similar to the one that desire theories had to face. That is: do we represent the distinctive unpleasant-feeling experience as bad because it is bad in itself, or is the distinctive unpleasant-feeling experience bad because we represent it to be bad? The distinctive feeling theory view seems to be taking the second horn of this dilemma. However, this is a dilemma. The first horn is much more intuitive, yet unavailable for distinctive feeling theory, if we want to account for the badness of the distinctive unpleasant feeling by appealing to the evaluative judgement about such feeling. The second horn is problematic since it leads an issue of justification. There seems to be no justifying reason, no good reason, to judge that the unpleasant distinctive feeling is bad, if there is nothing bad about it. I think that the best solution for the distinctive feeling theory is to take the first horn of the dilemma, and not to explain the badness of an unpleasant pain by appealing to evaluative beliefs.

This was precisely the appeal of this kind of theory in contrast to the desire theories discussed in the previous chapter. If unpleasantness is a feeling that explains and rationalises other mental states and actions, it is because the feeling is bad in itself. Being bad is an intrinsic property of the unpleasantness of pain. It is in virtue of this felt unpleasantness, that is bad in itself, that pain motivates and justifies actions. In fact, the idea that unpleasantness is bad in itself is widely accepted. Most content theories take unpleasantness to entail non-instrumental badness, even if they fail to explain why this is so, purely in terms of mental content. Desire theories also take unpleasantness to be intrinsically bad, and they explain this by appealing to desire frustration; although they

have to drop the strong intuition that unpleasantness is something felt in order to deal with the heterogeneity problem. I think that the solution for the distinctive feeling theory to be able to account for the normativity associated with the unpleasantness of pain is to establish that there are some phenomenal properties that are also normative. Unpleasantness is an example of this.

I think that the best strategy for a theory that claims that unpleasantness is a phenomenal property is to establish that such phenomenal property is bad in itself. In other words, to claim that the difference between unpleasantness and other phenomenal properties is that unpleasant experiences *feel bad*. Following Smuts, "to '*feel good'* is about as close to an experiential primitive as we can get… Phenomenology is likely our best tool… one cannot help but make recourse to metaphors. Here are a few: we might say that the pleasurable experience glows, we feel a warm feeling; the *good feeling* hums." (Smuts, 2011, p. 11, my emphasis). Bramble does not need to account for unpleasant pains being linked to motivation and normativity by appealing to normative beliefs, I think. Unpleasantness is an intrinsically motivational and normative phenomenal property because it entails badness. This is the starting point. We might not have given an explanation of why this is so, but there might not be an explanation for this. However, even if we grant that a distinctive unpleasant-feeling experience is bad in itself, and therefore it is a motivating and normative reason, there is still one crucial remaining issue. A theory such as Bramble's cannot account for the heterogeneity of unpleasant experiences.

3.2 Heterogeneity problem for the distinctive feeling theory

The heterogeneity of unpleasant experiences is a fundamental problem for the distinctive feeling theory. According to the heterogeneity problem, even if all unpleasant experiences feel unpleasant, there is no single unitary *phenomenal property*, no unitary feeling, in virtue of which all and only unpleasant experiences qualify as unpleasant. The distinctive feeling theory consists in claiming the opposite, that all and only unpleasant experiences qualify as unpleasant because they all share the same and only unitary distinctive feeling. According to the heterogeneity problem, if we make a list of all of the experiences that are unpleasant and we introspect them, we will not be able to find one single distinctive feeling that all and only these experiences have that could explain what makes all these experiences unpleasant. In other words: of course all and only unpleasant experiences share one property, they all are unpleasant! But the key intuition of the heterogeneity

problem is that all these unpleasant experiences are not unpleasant because they all share the same and only unitary phenomenal property of unpleasantness.

Bramble denies one of the intuitions behind the heterogeneity problem: even if it seems as if there is no unitary feeling in all and only unpleasant experiences, there is in fact such feeling. However, this feeling is not easy to notice because it is 'permeating' the sensory experiences. This permeation is meant to explain why, after simple introspection, it *seems* as if there was no distinctive phenomenal property that unifies all pleasant or unpleasant sensory experiences, respectively. In other words, the distinctive feeling is there, it is just that we cannot easily introspect it and become aware of it. In Bramble's words:

> Consider what 'the pleasant feeling' would have to be like if the distinctive feeling theory is to be at all plausible. It would have to be the sort of feeling that can occupy an experience, and so make it count as pleasant, by *permeating* it. Consider, for example, pleasant experiences of listening to Bach, eating a juicy peach, solving a puzzle, sunbathing, etc. Clearly, if 'the pleasant feeling' exists, it does not make these sort [*sic*] of experiences pleasant by being 'tacked on to them', so to speak, in any crude fashion. Instead, it must be the sort of feeling that can come in extremely low intensities, and very finely discriminable locations within one's experiential field, so that it can come scattered throughout one's experiential field. If the distinctive feeling theory is correct, and I enjoy listening to Bach, while you do not, then the difference between our experiences of Bach has got to be that mine is *permeated* by 'the pleasant feeling', while yours is not. (Bramble, 2013, p. 209–210)

Bramble only focuses on the heterogeneity of pleasant sensory experiences, but the same can be said for unpleasant sensory experiences and, more specifically, for unpleasant pains. That is to say, according to Bramble all unpleasant pains share the same phenomenal property of feeling unpleasant, but we cannot identify this shared feeling because it permeates the pain experiences, as opposed to being roughly 'tacked on'. If this were right, what does it mean for a distinctive feeling to *permeate* our experiences, instead of it merely being *tacked on* to them? I think the analogy with colour experience will be helpful to make sense of Bramble's proposal.

Let us suppose, as I have said before, that the distinctive unpleasant-feeling is analogous to the distinctive red-feeling that typical experiences of seeing something red have. According to the distinctive feeling theory, there is a shared distinctive unpleasantness

when we experience a cut, a burn, a headache, etc. Although, as Bramble says, the distinctive feeling of unpleasantness is hard to find after introspection, given the subtle permeation with which the unpleasantness modifies the pain experiences. Similarly, one recognises some similarity in the experiences of seeing a ladybird, of seeing a stop sign on a traffic light, and of seeing an apple. All of these experiences are phenomenologically similar because they have some shared red-feeling-ness. If this is a fair analogy, which I think it is, we can try to understand what it would mean for distinctive red-feeling-ness to be permeating our visual experiences, so we can make sense of the permeation of the distinctive unpleasantness that Bramble brings to the table.

Suppose that you have the visual experience of a Magritte painting. How could red-feeling-ness permeate this visual experience? I can think of two ways of making sense of the notion of permeation. First, in the way that water *permeates* a sponge by filling its cavities with water, if the Magritte painting were filled with tiny red dots, these tiny red dots would be dispersed and filling the cavities of the painting, as it were. These subtle red dots in the painting translate into red-feeling-ness permeating your experience of a Magritte painting. In the same way that the dots are 'scattered throughout one's experiential field', red-feeling-ness is spread throughout your visual experience. The dots still allow you to identify what you see as a Magritte. Similarly, the permeating red-feeling-ness still allows you to identify your visual experience as one of a Magritte painting. To get an idea of this, see figure 1 below. This is one possible way to try to make sense of how a distinctive feeling can permeate an experience.

An alternative is to understand the permeation as a filter. It could be that one's visual experience is permeated in the sense that it is being modified as if it had a filter. Imagine you go to the museum wearing red shaded glasses, your experience of a Magritte painting is then filtered with a red tonality, you still identify the painting as a Magritte but now it has a red permeation. Similarly, your visual experience of the painting has now some red-feeling-ness permeating it. See figure 2 to get an idea of how this would be. Even if figure 2 might not be exactly what happens when one wears coloured shades, the modification of the pixels of the image turned closer to the red spectrum is meant to exemplify how your visual experience of this image now has some red-feeling-ness, that it would not have had without modifying the colour of the image. This image modification results, the idea goes, in your visual experience having permeating red-feeling-ness. Sometimes the modification of the painting might be so subtle, so fine grained, that it would result in an almost unnoticeable red-feeling-ness at an experiential level.

Now, contrast these two possibilities as opposed to what being 'tacked on' would mean in terms of a phenomenal property of experience. If someone attached a big red stripe to the Magritte painting, the red stripe would be literally tacked on to the painting. This would not translate into your experience being permeated by red-feeling-ness, but rather into having red-feeling-ness being tacked onto it. To illustrate this, see figure 3 below. The unpleasant feeling is not 'tacked on' to the pain experience, Bramble says. The feeling can come "in extremely low intensities and very finely discriminable locations within one's experiential field". In other words, the fact that our unpleasant pain experiences are permeated with such distinctive unpleasantness does not imply that this unpleasant feeling will always be easily localisable or discernible. So, what Bramble proposes should be rather similar to figures 1 or 2 and not at all like figure 3. I think this gives us a better idea of what Bramble could mean when he says that a distinctive feeling permeates our pain experiences.

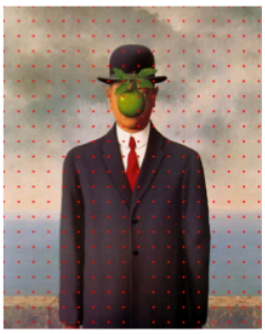

Fig. 1

Fig. 2

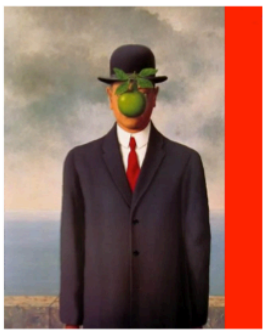

Fig. 3

However, I think Bramble's appeal to permeation is unsatisfying when it comes to dealing with the heterogeneity of unpleasant experiences. In short: there is something quite odd about something being *distinctive* yet going *unnoticed*. As I said earlier, if something is distinctive, this can be understood as meaning that it is noticeable. That is precisely one of the initial appeals of the distinctive feeling theory. Namely, unpleasantness is a feeling and such feeling is available to introspection. The distinctive unpleasant-feeling should be at least noticeable after careful introspection, but it isn't. Furthermore, careful introspection shows us that there is no shared unitary unpleasant feeling among all and only unpleasant experiences. Consider a headache, a stomach ache, the feeling of a burn, the pain produced by a cut. Do all of these have one single shared distinctive unitary unpleasant-feeling? Now consider the feeling of being nauseous, having a cramp, an itch, or numbness. Do all of these have one single shared distinctive unpleasant-feeling that is also shared with all unpleasant pains? It does not seem to be so. Bramble's proposal about the distinctive feeling is problematic in various ways.

First, if by *distinctive* Bramble means *easily noticeable*, the distinctive feeling is not easily noticeable. Once Bramble introduces the notion of permeation, he admits that it is not very easy to introspect an unpleasant distinctive feeling. Bramble seems to argue that the distinctive feeling is almost unnoticeable, since it is permeating the experience, but how can something almost unnoticeable be distinctive? If this is not contradictory, it is at least very odd. Second, if by *distinctive* he merely means *noticeable*, it is not very obvious that we can notice a distinctive unpleasant feeling in a pain experience, in the way that we can notice red-feeling-ness in visual experiences. Whereas we seem to be able to introspect a phenomenal quality of red-feeling-ness in various visual experiences, this does not seem to apply so clearly to the unpleasant-feeling. Third, even if we accept that we can introspect an unpleasant-feeling in at least some unpleasant experiences, e.g., in various pains, there is no unitary distinctive feeling that is present in all and only the experiences that we introspect as being unpleasant. This last intuition is key for heterogeneity being a problem for the distinctive feeling theory.

If the unpleasant distinctive feeling is a phenomenal property, it should be available to introspection, even if it is permeating the sensory experiences. This introspection should allow us to identify all unpleasant experiences as sharing the same unitary unpleasant feeling, a common phenomenology. The distinctive feeling theory proposes that there is one unitary phenomenal aspect in virtue of which all and only unpleasant experiences qualify as unpleasant. When we cannot find such feeling, Bramble argues that such feeling is hard to find because it is permeating the experience. However, even if it is hard to notice, it should at least be noticeable after careful introspection. Careful introspection does not show any unitary unpleasant phenomenal commonality among all and only unpleasant experiences. Bramble's solution to the heterogeneity problem relies heavily on the rather metaphorical notion of permeation. However, even when we try to make sense of such a notion, as we do by comparing unpleasantness with red-feeling-ness, we cannot identify one unitary feeling that all and only unpleasant experiences have. The heterogeneity of unpleasant experiences shows that the distinctive feeling theory is wrong.

3.2.1 A possible solution for the distinctive feeling theory?

I think that there might be another way to try to show that there must be some unitary phenomenal feature that is shared among all and only unpleasant experiences. The mere fact that we can compare different unpleasant experiences among them entails that they all share a qualitative aspect of unpleasantness.[17] So, the idea goes, if you compare items *A* and *B*, where both are unpleasant and you distinguish which is more unpleasant to you, this should mean that *A* and *B* share a qualitative property in virtue of which you evaluate which one is more unpleasant. You can then do the same process between *B* and *C*, *C* and *D*, and so on. At the end, you will have a list of unpleasant experiences that all share a phenomenal property of unpleasantness, which should show that the heterogeneity of unpleasant experiences is not actually a problem for an approach such as the distinctive feeling theory. The fact that we can compare different unpleasant experiences, in virtue of being unpleasant, must show that they all have the same type of unpleasantness in common, the same and only unpleasant feeling. If this is correct, the supporter of the distinctive feeling view could endorse such an argument. However, I will show that this argument based on the comparison of various unpleasant experiences cannot really dissolve the heterogeneity problem.

There are various reasons why this comparison argument is problematic. The first problem is that even if we *claim to be judging* between two different experiences on the basis of their unpleasantness, this does not imply that we are actually judging the experiences based on their unpleasantness as understood as a phenomenal property, some feeling that is introspectively available. Instead, it could be that what we are really doing is judging them based on other mental states. If you were asked 'which is more unpleasant to you *A* or *B?'*, and you answered '*A'*, this answer does not entail that you are making a judgment based on a phenomenal shared feature between *A* and *B*. It could be that you are reporting which one you *like* more, which one you *desire* more, which one *causes* less suffering to you, etc. In either case, the basis of the judgment is not necessarily a shared qualitative unpleasant feeling that *A* and *B* have in themselves.

Let me try to clarify this with an analogy. For example, you could be asked 'what is more pleasant for you: a) hiking in the mountains, or b) lying on the beach?' The fact that you have a preference for any of these options does not imply that your judgement is based on any shared phenomenal property common to these two possible scenarios. You might prefer the beach because it reminds you of your childhood vacations, for instance, which

[17] I owe this idea to Jennifer Corns.

is not something that the experience of being on the beach entails, and neither something that lying on the beach or hiking in the mountains necessarily share. The preference that you might have for one option over another does not entail a comparison in virtue of a mutual property that both items have. In the same way, when we compare different experiences, regarding which one is more pleasant or unpleasant, our judgment is not necessarily based upon a phenomenal property of these experiences.

There is a second, more fundamental problem for this comparison argument. Even if we compare different experiences because they share a phenomenal property, this does not entail that the heterogeneity of these experience is not still a problem for the distinctive feeling theory. The defenders of this theory could claim that making a list of unpleasant experiences does entail that all the items in the list are unified by one single shared phenomenal property. They might think this because, otherwise, it would have been impossible to make the list. In a sense this is correct. The fact that we come up with a list of items that are being compared because they are unpleasant implies that all the items have a shared property: being unpleasant. However, this list does not entail that all the items share the same way of being unpleasant, that is, that there is a unitary feeling that all and only unpleasant experiences share and by dint of which they are unpleasant. Even if it were true that items *A* and *B* are unpleasant and that *B* and *C* are also unpleasant, this does not necessarily imply that *A*, *B,* and *C* share the exact same type of phenomenal unpleasantness, that is, that they are all unpleasant *in the same way*. *B* may share some phenomenal property with *A* and with *C*, but it is not necessary that *A* and *C* also share the exact same phenomenal property between them.

Let us use an analogy to clarify this point. First, you are asked to judge what is more *delicious*: a) a sandwich with tomatoes, or b) a soup with tomatoes. Then, you are asked to judge what is more delicious: b) that soup with tomatoes, or c) a carrot soup. The sandwich, the tomato soup, and the carrot soup are all delicious. However, this does not mean that they are delicious on the exact same grounds. On the one hand, the tomato sandwich and the tomato soup may be judged in virtue of the quality of their tomatoes. On the other hand, the tomato soup and the carrot soup might be judged in virtue of their consistency. This three-item-comparison *does* imply that all compared items are delicious, but they do not have to be delicious in the same way. Some might be delicious *qua* their tomato-ness and others *qua* their consistency. In this case, being delicious is not one single qualitative property that is shared among the three items; deliciousness is not a single shared ingredient that is present among all instances. When we are comparing different

items in terms of their deliciousness, we might not be comparing them *qua* the same type of deliciousness.

We can say something similar for the comparison of unpleasant experiences. You might compare if it is more unpleasant for you: a) to feel nauseous, b) to have a headache, or c) to feel an itch. The comparison is possible, and at the end you would have a list of three unpleasant experiences. These comparisons could be made with many more experiences and you could create a quite large list of unpleasant experiences. However, this does not entail that all the items in the list share the same property of unpleasantness, if it is understood as a unitary qualitative feeling that can only have variations in terms of intensity. The comparison argument can show that all items are unpleasant, but it *does not* imply that all of its items are being compared on the grounds of being unpleasant in the same phenomenal way. In other words, this comparison does not entail that the feeling of nausea, of a headache, and of an itch all have the exact same phenomenal property, understood as an introspectively available unitary feeling that comes in varying intensities.

The question that a defender of the distinctive feeling theory should ask in relation to the heterogeneity of experiences is: which is more unpleasant regarding the *exact same* shared felt qualitative feeling of unpleasantness *A* or *B*, *B* or *C*, *C* or *D*? If we were able to answer this question with no difficulties, then we could show with no doubt that heterogeneity of unpleasant experiences is not a problem for the distinctive feeling theory. However, we cannot answer this question effectively. This is precisely what the heterogeneity problems shows, that is, even if we accept that many diverse experiences are all unpleasant, this unity does not seem to be explained in virtue of a unitary shared phenomenal property of unpleasantness, of one and only distinctive feeling.

3.3 Conclusion

I think that the distinctive feeling theory is useful to get a grasp of an approach for which unpleasantness is a phenomenal property that is not taken to be reducible. Even if Bramble does not argue that this phenomenal property is bad in itself, I think that this is the best strategy for a theory such as this to account for the unpleasantness of pain being a motivating and a normative reason. We can explain that we take a painkiller because the unpleasantness of pain is bad in itself. The unpleasantness of pain being bad renders our actions intelligible, it explains why an unpleasant pain is a motivating reason. The badness of unpleasantness also explains why taking a painkiller is desirable, why it is good to stop

feeling something unpleasant. The intrinsic badness of unpleasantness accounts for an unpleasant pain also being a normative reason.

However, even if we accept that this phenomenal property is bad in itself, the distinctive feeling theory fails to deal with the heterogeneity of unpleasant experiences. In the upcoming chapter I will consider two more theories that account for the unpleasantness of pain in terms of a phenomenal property. In contrast to the distinctive feeling view, these other theories can appeal to phenomenal variations in order to deal with the heterogeneity of unpleasant experiences. I will show how these theories rely on the notion of *dimension* in order to deal with the heterogeneity problem. Moreover, I will explain how one of these views successfully provides a solution to the heterogeneity problem.

CHAPTER 4: THE DIMENSIONAL THEORIES OF UNPLEASANTNESS

4.0 Introduction

In this chapter I will explain other theories that take unpleasantness as a phenomenal property. These accounts, in contrast to the distinctive feeling theory, add the notion of *dimension* to unpleasantness. I will focus on two theories that opt for this strategy. I will evaluate how adding the notion of dimension to unpleasantness can be used to deal with the heterogeneity of unpleasant experiences. I will show that the notion of dimension, understood in line with the determinable-determinate distinction, is useful to argue for unpleasantness as a phenomenal property, while also solving the heterogeneity problem.

According to the first theory, the *hedonic dimension theory*, unpleasantness is a dimension along which sensory experiences may vary. The idea comes from Kagan (1992), even if he is generally concerned with pleasant experiences. In the case of pain, for instance, unpleasantness can be explained as the felt dimension along which sensory pains can vary. Broadly, according to this proposal, unpleasantness is to pain what volume is to sound. This dimension is meant to account for the variations among different unpleasant experiences, e.g., this dimension can be used to explain the different intensities with which unpleasant pains might be felt. I will divide the discussion of this theory into two parts.

First, I will briefly consider a critique from Bramble (2013). Bramble thinks that Kagan's analogy with volume is inapt: whereas volume is an essential property of sound, being hedonic is not an essential property of sensory experiences. I will show that there is a way to deal with Bramble's critique. Second, I will show that even if the analogy between unpleasantness and volume can be apt, Kagan's suggestion to take unpleasantness as a dimension is not enough to give an answer to the heterogeneity problem. Even if the qualitative dimension approach is useful to account for a type of variation among unpleasant experiences, i.e., to account for different intensities of unpleasantness, it does not account for other important forms of phenomenal variation among unpleasant experiences.

Given that the hedonic dimension theory is not able to solve the heterogeneity problem, I will consider a different theory that gives an answer to this problem by appealing to the determinable-determinate distinction (Crisp, 2006). First, I will explain Crisp's proposal to solve the heterogeneity problem based on the determinable-determinate distinction. I will explain how we should understand Crisp's account in order to give a straightforward

answer to the heterogeneity problem. Broadly, based on this proposal, we can understand that in the same way that scarlet is a way of being red, and red is a way of being coloured, a headache is a way of being a pain, and pain is a way of being an unpleasant experience: being a *pain* is a determinate of the determinable being *unpleasant*.

However, this theory has a fundamental difficulty: accounting for cases of pain not being unpleasant. The account derived from Crisp's proposal about pain and unpleasantness entails that pain is unpleasant. In this view it is necessary that pain is unpleasant, as it is necessary that scarlet is red. It is impossible for a pain not to be unpleasant. However, there are various cases that suggest that pain might not be unpleasant. In the upcoming chapters, I will revise these cases.

4.1 Hedonic dimension theory

The hedonic dimension theory proposes that unpleasantness should be understood as a dimension along which experiences such as pain may vary. This proposal relies on an analogy: volume is to sound what pleasantness or unpleasantness are to sensory mental states. If this analogy is correct, one of the virtues of this proposal is that it can capture an important feature of unpleasant experiences, i.e., that unpleasantness can have different intensities. There are some experiences that are more, or less, unpleasant than others. In the same way that volume is a property of sound, and that this property may come with different intensities, being unpleasant is also a property of certain sensory mental states and this property can also have different intensities. In Kagan's words:

> An alternative move is to identify pleasantness not as a component of experiences, but rather as a *dimension* along which experiences can vary. As an analogy, consider the loudness of auditory experiences — that is, sounds. It is obvious that loudness or volume is not a *kind* of sound. And it seems plausible to insist that loudness is not a single kind of component of auditory experiences. Rather, volume is a dimension along which sounds can vary. It is an aspect of sounds, with regard to which they can be ranked… Similarly, then, pleasure might well be a distinct dimension of mental states, with regard to which they can be ranked as well… For it seems to me that there is a sense in which a specific volume is indeed an ingredient of a given sound, along with a particular pitch, and so forth. (Similarly, intensity or saturation is an ingredient of colors, along with hue.) (Kagan, 1992, p. 172–173)

Let us try to make sense of Kagan's idea. His proposal is about pleasantness, but something very similar can be said for unpleasantness, that is, unpleasantness is an aspect along which different experiences can vary, including pain. Since volume is a dimensional property with intensities, we can explain that some sounds *are louder* than others. Similarly, if unpleasantness is a dimensional property with intensities, we can explain that some unpleasant experiences are more intense. Moreover, if we take unpleasantness to be bad in itself, we can also explain that some unpleasant experiences are worse, they *feel worse*. I think that the key feature of Kagan's idea is that if volume is a dimension of sound, this means that volume can have different intensities, and, therefore, that sound can also have different intensities in terms of its volume. Similarly, if unpleasantness is like volume, this means that a sensory experience can have different intensities in terms of how bad it feels.

4.1.1 The essentiality of unpleasantness

With this in mind, let us now analyse some of the difficulties that the hedonic dimension theory may encounter. First, Bramble (2013) argues that Kagan's analogy is inapt. "Consider that, for most pleasant experiences, one can reduce their pleasantness to nothing, while leaving the experience intact, whereas one cannot ever reduce the volume of an auditory experience to nothing and still be left with the auditory experience in question." (Bramble, 2013, p. 209) I think that a simple way of explaining Bramble's critique is to focus on volume as a property of sound. Whereas volume is an *essential property* of sound, being pleasant or unpleasant is not essential to mental states such as sensory experiences. Sound is often understood as a physical phenomenon, as a perturbation in a medium such as air; these perturbations can be measured in waves and the standard way of understanding sound's volume is in terms of the amplitude of the waves. The bigger the volume the bigger the wave. If the wave has no size, this means that there is no perturbation in the medium, that is, that there is no sound. This is why volume is essential to sound. In contrast, being pleasant or unpleasant is not essential for sensory experiences. For example, we could have a taste experience of drinking coffee without it being pleasant or unpleasant. Whereas sound cannot exist without volume, certain sensory experiences can exist without having a hedonic dimension. This, it is claimed, shows that the analogy is inapt.

I think that there are two key things to mention about Bramble's critique. First, it is important to have in mind that the heterogeneity problem is not concerned with

unpleasantness being essential or not to sensory experiences. That is to say, the intuition behind the heterogeneity problem is that there is no unitary phenomenal feeling in virtue of which all and only unpleasant experiences count as being unpleasant. This intuition can be addressed regardless of unpleasantness being essential or not to sensory experience, and I will explain in detail how Kagan's suggestion about unpleasantness being a dimension can be used to try to deal with the heterogeneity problem in the upcoming subsection. The second important thing to mention about Bramble's critique is that there is a simple way of dealing with it.

I think that there is a simple and elegant way to take Kagan's analogy between the unpleasantness of experiences and the volume of sounds. We could take his analogy as comparing the unpleasantness of *unpleasant experiences* with sound's volume. That is, if volume is essential to sound, we can argue that being unpleasant is essential to unpleasant mental experiences. I think that Bramble's mistake consists in understanding Kagan's analogy as referring to hedonically neutral experiences. Rather, we should understand that in the same way that volume is an essential property of sound, unpleasantness is an essential property of unpleasant experiences. In this way, even if it is true that a sound must have volume in order to exist, the analogy is apt because an unpleasant experience has to be unpleasant in order to exist as an unpleasant experience. An unpleasant pain has to be unpleasant in order to count as an unpleasant experience. The analogy is finally apt regarding essentiality.

4.1.2 The heterogeneity problem for the hedonic dimension theory

There is, however, another much more important problem for the hedonic dimension theory: it cannot account for the heterogeneity of unpleasant experiences. Let me show this. Volume is a property that all and only sounds have and unpleasantness is a property that all and only unpleasant experience have. The variations that volume provides in terms of sound can explain some form of intensity of sound. Sound can have various intensities regarding volume. I think that if we understand Kagan's analogy of volume, this is helpful to make sense of how unpleasant experiences can also have various intensities. The intuition behind the heterogeneity problem is that there is a phenomenal variation among unpleasant experiences that a unitary unpleasant feeling cannot account for. We could then try to account for such phenomenal diversity of unpleasant experience in terms of the hedonic dimension theory.

> Recognition of the qualitative differences between the sounds of a symphony, rain falling, and a bird chirping, does nothing at all to call into question our ability to identify a single dimension —volume —with regard to which these and other sounds can be ranked… Recognition of the qualitative differences between the experiences of hiking, listening to music, and reading philosophy, need not call into question our ability to identify a single dimension —pleasure — along which they vary in magnitude. (Kagan, 1992, p. 172–173)

This is not problematic so far. We can agree with Kagan about the diversity of pleasant experiences and the same can be said for unpleasant experiences. The problem with the hedonic dimension theory is that unpleasantness varies in ways that are not only explained in terms of intensity. Whereas volume *only* varies in terms of intensity, unpleasantness *not only* varies in terms of intensity. The hedonic dimension theory has to show that all and only the experiences that are unpleasant are unified in virtue of being unpleasant, in virtue of having the same phenomenal unity, the same phenomenal dimension, and that the only variation among all these unpleasant experiences, qua their unpleasantness, is regarding the intensity of the same unitary felt unpleasantness. The hedonic dimension theory is not able to meet this intuition, and this is why the heterogeneity of unpleasant experiences is a problem for this theory. This is a fundamental problem for this account.

If we think that it is the same qualitative dimension that unifies all unpleasant experiences, we must show that the same and only phenomenal dimension can account for all the phenomenal variations among unpleasant experiences. If we understand unpleasantness as analogous to volume, this means that all the variations of unpleasantness can be reduced to variations in intensity. However, unpleasant experiences vary in ways that are not only accountable in terms of intensity. I agree that hiking, listening to music, and reading philosophy can all be pleasant, but I do not think that their pleasantness can only vary in terms of intensity. The same can be said for unpleasant experiences, i.e., it is not particularly hard to accept that experiencing grief, nausea, suffering from severe burns, having a headache, etc. are all unpleasant. What is controversial is that all these experiences are all unpleasant in virtue of the same shared phenomenal aspect that all and only these experiences have, and that the only difference regarding their unpleasantness is that some of these experiences are more unpleasant than others. Kagan admits that it is odd to say that there is something phenomenal shared among all these experiences, but he insists that:

> Once we have a picture like this in mind, we might in fact be prepared to insist that there *is* a sense in which pleasure is an ingredient common to all pleasant experiences… Thus, pleasantness might well be considered an ingredient of (conscious) mental states in general, albeit an ingredient that we will only notice if we "chop up" experiences in some nonstandard ways. But whether or not pleasure can be helpfully viewed in this way as an ingredient of experiences, the possibility remains that it is a single, specific dimension along which experiences vary. (Kagan, 1992, p. 172–173)

Kagan seems to suggest that we will find the shared qualitative property, the shared phenomenal dimension, once we have "chopped up the experiences in some nonstandard ways". I am not sure what this could mean, but I do think that the crucial point about Kagan's view is that there is a "dimension along which experiences vary". This idea of dimension allows us to shed light on the phenomenal heterogeneity of hedonic experiences. However, if we understand unpleasantness as being similar to volume, I think that the analogy is limited because unpleasantness varies in ways that volume does not. Whereas volume varies in intensity, which can be understood in terms of a wave's size, unpleasantness *not only* varies in intensity. If we take seriously the analogy between sound's volume and the unpleasantness of unpleasant experiences, we will find that the analogy is inapt. It is not inapt regarding the essentiality of these properties. The analogy is inapt because the variations of unpleasantness cannot be captured in terms of the variations of volume.

Let me clarify this last idea. For argument's sake, let us understand sound as being constituted of three distinct essential dimensions: duration, i.e., for how long a disturbance occurs in a medium; volume, which can be understood in terms of the size of the waves that represent the strength of sound, to put it simply; and pitch, which is the frequency of the waves within the duration of sound. If sound's volume is an essential dimension of sound, this implies a few things, for example: i) two different sounds may vary only *qua* their volume, i.e., both sounds have the same duration and the same pitch, but the intensity of their volume is different, and ii) two sounds may have the exact same intensity of volume but vary regarding to their other essential properties, that is, vary in their duration or their pitch.

In contrast, when we try to do something similar for unpleasant experiences, we notice that the heterogeneity of unpleasantness cannot be explained in terms of the hedonic dimension theory. We cannot give a similar account for unpleasantness, where the

intensity of unpleasantness is the only kind of variation among all and only unpleasant experiences. To show this, I will first consider unpleasant pains. Suppose that all unpleasant pains are mental states that are constituted by these essential properties: felt duration, felt location, a pain sensory aspect, and felt unpleasantness to a certain degree. Consider two headaches; I take it that we could have two different pains that vary only in the way that they are unpleasant. That is, we could have two headaches, a dull one and a sharp one, that have the same felt duration, felt location, a headache-like feeling, and the same intensity of unpleasantness, i.e., they feel equally bad, and yet they are different types of headache experiences.

If this is correct, this shows that there are phenomenal variations that we cannot explain merely in terms of the intensity of how bad the experiences feel. According to the hedonic dimension theory, intensity is the only type of variation among unpleasant experiences in terms of their unpleasantness. However, what I am trying to show is that different unpleasant pains can vary in the way in which they are unpleasant, without being different in terms of how bad they feel. When two sounds vary in volume, their difference *qua* volume can only be a difference in terms of intensity. In contrast, when it comes to unpleasantness, the differences between two unpleasant experiences *qua* their unpleasantness are not only differences regarding their intensity, but also differences regarding other ways of being unpleasant.

If this is not persuasive enough, let me offer a few more cases. The heterogeneity of unpleasantness becomes clearer the broader the type of unpleasant experiences we want to explain. Compare a migraine with a stomach ache or a toothache. They vary in location, clearly, but are they all *achy* because of the exact same phenomenal quality dimension? Suppose that they all rank with the same intensity *qua* their unpleasantness and that they all last the same amount of time. There is going to be an obvious difference regarding location, there might also be a difference regarding a sensory component of the experience, but it does not seem clear to me that they are all unpleasant in the same phenomenal way.

Compare a migraine with feeling nauseous. Their unpleasantness may differ not just in intensity, like the volume of two sounds, but in other ways too. Contrast an intense migraine with intense grief (i.e., with the feeling that one can gets when someone dear dies). Are these two intense unpleasant experiences unpleasant in the exact same phenomenal way? I think that the careful comparison between all these unpleasant experiences shows that there is no unitary feeling among all and only these experiences

and in virtue of which they all qualify as unpleasant. I consider that Kagan's notion of dimension offers some improvement over the distinctive feeling theory, since with it we can account for a type of variability within unpleasant experiences. The hedonic dimension theory acknowledges that unpleasant experiences may have different intensities. However, this is not enough to shed light on all the possible variations regarding the unpleasantness of unpleasant experiences.[18]

In conclusion, the heterogeneity of unpleasant experiences is an issue for the hedonic dimension theory. Even if the notion of unpleasantness as dimension serves to give us a better grasp of the variations among unpleasant experiences, this theory does not seem to capture the heterogeneity of unpleasant experiences. If we take unpleasantness to be phenomenal and bad in itself, we can understand the intensity of unpleasant experiences in terms of how bad they feel. However, how bad an unpleasant experience feels does not seem to be the only way unpleasant experiences may vary in terms of being unpleasant. In the next section I will consider a different theory that accounts for the unpleasantness of pain in terms of a phenomenal property and that is able, I will argue, to solve the heterogeneity problem.

4.2 The solution to the heterogeneity problem

The heterogeneity of unpleasant experience is a major issue for theories that take unpleasantness to be a phenomenal property: there is a strong intuition that there is no unitary phenomenal unpleasant aspect that is present among all and only unpleasant experiences, and in virtue of which all these unpleasant experiences qualify as unpleasant. In other words, even if we accept some phenomenal unity among all unpleasant experiences, it seems that we cannot account for this unity by appealing to a phenomenal property, the same unpleasant feeling that all and only these experiences have. Something very similar can be argued for the variety of pleasant experiences. So how can we preserve the intuition that unpleasantness is phenomenal, while explaining the wide phenomenal diversity within unpleasant experiences?

---

[18] Aydede (2014) proposes a similar approach to Kagan's that can be used to account for different intensities of unpleasantness. "Just the fastness or slowness of dances can be recognized across all different types of dances, the pleasantness or unpleasantness common to various otherwise quite different sensations is detectable, indeed introspectively available." (Aydede, 2014, p. 113) However, I think that this proposal faces similar worries since the intensity of unpleasantness is not the only type of phenomenal variation among unpleasant experiences.

In order to deal with this problem, while maintaining the intuition that unpleasantness is something that we somehow feel, it has been proposed that we should explain the way in which experiences can be pleasant or unpleasant by appealing to the determinable-determinate distinction. Crisp (2006) proposes this with respect to pleasure:

> If the advocate of heterogeneity is seeking in enjoyable experiences something like a special sensation, such as sweetness, or a tingle or feeling located in a certain part of the body, such as an itch or pins and needles, or indeed something like a perceptual quality such as redness, she will fail. But there *is* a way that enjoyable experiences feel: they feel enjoyable. That is, there is something that it is like to be experiencing enjoyment, in the same way that there is something that it is like to be having an experience of colour…Enjoyment, then, is best understood using the determinable-determinate distinction, and the mistake in the heterogeneity argument is that it considers only determinates. Enjoyable experiences do differ from one another… But there is a certain common quality… feeling good... The determinable–determinate distinction also helps us to be clear about the role of 'feeling' in this analysis: feeling good as a determinable is not any particular kind of determinate feeling. (Crisp, 2006, p. 109)

There are a few rather terminological things to say about Crisp's passage. First, Crisp's idea can be equally applied to experiences that *feel bad*, i.e., several instances of experiences can be explained as determinates of the common feeling-bad determinable. Instead of referring to these determinables as 'feeling good' or 'feeling bad', I will refer to them as being pleasant or being unpleasant. The change in the way in which I will refer to these determinables is primarily terminological, in order to maintain coherence with the previous chapters. The common thing between feeling nauseous, feeling itchy, and the experience of a headache, is that they are all unpleasant.

Second, I should also point out that when I talk about determinables and determinates, I will be referring to properties. That is, as in previous chapters, being unpleasant is a property of mental states, and I will talk about different ways of being unpleasant that are also properties of mental states. I will show that these properties of mental states should be understood based on the determinable-determinate distinction, if we want to solve the heterogeneity problem. Moreover, I will often compare the property of being unpleasant, or unpleasantness, to other properties. For instance, I will compare unpleasantness to the

property of being red, or redness, when it comes to objects, or to the property of being red-feeling, or red-feeling-ness, when it comes to visual experiences.

Now that these terminological issues are established, we can start explaining how the determinable-determinate distinction will serve us to deal with the heterogeneity of unpleasant experiences. I propose that a good way of making sense of this distinction is to say that different determinates vary among themselves *non-additively*. I propose that this understanding of the determinable-determinate distinction is crucial if we are to make sense of this proposal. I will first exemplify this notion by considering the property of being coloured as a determinable. To say that being coloured is a determinable is to say that the colour properties that fall under it differ from one another *non-additively*. In other words, the difference between being red and being blue is *not* that being red consists in being coloured plus being *X*, whereas being blue consists in being coloured plus being *Y*, but rather that being red consists in being coloured in a particular way, and being blue consists in being coloured in a different particular way. Moreover, being coloured is the sum of the different ways of being coloured; being coloured is nothing more that being either red, or blue, or yellow, or green, etc. Some determinates can also be determinables of other determinates. For example, being red can also be a determinable with multiple determinates such as being magenta, crimson, scarlet, etc. The same can be said for blue, for example being baby-blue or navy-blue can be determinates of the blue determinable.[19]

By means of contrast, let us consider other cases where the instances of a common group are different *additively*. Consider the property of being a young animal. All of the items that fall under this property share a distinctive ingredient, i.e., being young, *plus* something else. For example, being a kitten means to be young plus be a cat; being a cub is to be young plus be a bear; being a lamb means to be young plus be a sheep; and so on. That is to say, the properties of being a kitten, a cub, and a lamb are different ways of being a young animal, and they differ from one another in an additive way; they are not determinates of a common determinable. They are ways of being a young animal, but these different ways of being are not captured by the determinable-determinate distinction.

The key is to understand that different determinates of the same determinable are not unified by dint of sharing a distinctive ingredient plus something else. In this way, we can also apply the determinable-determinate distinction to properties of mental states. Something similar can be said for unpleasant experiences. Following this distinction,

[19] For more on the determinable-determinate distinction see Funkhouser (2006, 2014), Johnson (1921), Prior (1949), and Wilson (2017, 2009).

being unpleasant is not a common ingredient among the different unpleasant experiences. Instead, being unpleasant is a determinable with determinates such as being the feeling of nausea, of itch, of pain, etc. These different determinates are different ways of being unpleasant and they vary from each other non-additively. We can give an answer to the heterogeneity problem once we adopt this approach.

I think that a clear way of making sense of this is to make an analogy with being coloured. All ways of being coloured are different, yet they belong to the same kind. There seems to be no unitary coloured-ness that is present in being red, being blue, being yellow, etc. How can being red and being blue be different and yet both be ways of being coloured? The answer is that they are both determinates of the same determinable. Being red and being blue are different non-additively. They do not share a distinctive coloured-ness plus something else. The same can be said about the difference between being magenta and being scarlet; they do not share some unitary redness plus something else, they are different ways of being red non-additively. The details about what accounts for being red and being blue to be different, and for being magenta and being scarlet to also be different, are still missing. I will come back to this. However, what is important at this point is to understand that different determinates of a common determinable are different in a non-additive way. That is, they do not belong to the same kind by sharing a unitary ingredient plus something else.

Let us now apply this notion to different ways of being an unpleasant experience, in order to address the heterogeneity problem. If we try to make sense of Crisp's idea, this means that being unpleasant is a determinable with various determinates such as being an itch, being a pain, being a cramp, etc. This means that being an itch and being a pain do not share a common ingredient of unpleasantness plus something else that makes them different. They are different ways of being unpleasant non-additively. In the same way that there is no unitary coloured-ness shared between being red and being blue, there is no unitary shared unpleasantness between being an itch and being a pain. If we understand Crisp's proposal in these terms, the heterogeneity problem is solved, I claim. According to the heterogeneity problem there is no single unitary phenomenal ingredient common to all and only unpleasant experiences and by dint of which these experience are unpleasant. This is correct according to the non-additive understanding of unpleasantness as a determinable. All it means to have an unpleasant experience is to have a mental state with the determinate property of either being an itch, being a pain, being a cramp, etc. These different determinate properties entail unpleasantness without unpleasantness being a unitary ingredient common to all of them.

If we understand unpleasantness as a determinable, together with the non-additive understanding that I offer for the determinable-determinate distinction, the heterogeneity problem dissolves. Being a pain or being an itch are ways of being unpleasant that differ non-additively: being an itch and being a pain do not share some unitary unpleasantness. We can maintain that being unpleasant is a phenomenal property that is bad in itself, and also accept that being unpleasant is not a common unitary ingredient common to all and only unpleasant experiences.

4.2.1 The dimensions of unpleasantness

I think that even if the determinable-determinate distinction helps us to deal with the heterogeneity of unpleasant experiences, there is still an important element missing: we need to explain what makes different determinates vary from one another. What makes being red different from being blue? What makes being scarlet different from being magenta? In order to explain this, I think we need to focus again on the notion of *dimension*. Kagan was right in noticing that unpleasant experiences may vary in a way that is similar to the way in which sounds may vary. However, understanding the unpleasantness of experience as being similar to the volume of sound was limited to accounting for variations amongst unpleasant experiences other than their intensity. In this section I will develop Crisp's proposal in order to account for how different determinates of a common determinable vary non-additively. The differences among determinates of a common determinable can be explained in terms of variations along the multiple dimensions that might constitute a determinable. I think that Johnson's mention of colour as a determinable can be helpful at this point.

> [I]n fact, the several colours are put into the same group and given the same name colour, *not* on the ground of any partial agreement, but on the ground of the *special kind of difference* which distinguishes one colour from another; whereas no such difference exists between a colour and a shape. (Johnson, 1921, p. 176, my emphasis)

Similarly, Prior (1949, p. 13) says "redness, blueness, etc., all characterise objects, as we say, 'in respect of their colour'… And this is surely fundamental to the notion of being a determinate under a determinable". So what is this *special kind of difference* that. Johnson talks about? How is it that redness and blueness vary *in respect to their colour*, as Prior mentions? I think that for this we need to appeal once more to the notion of dimension.

This is key to understand how different determinates of a common determinable can vary from one another. The dimensions are the essential properties of determinable and in virtue of which each of its determinate varies from the rest. As I said, this is what was missing. We needed to know what makes different determinates of a common determinable different among them and this is the answer: it is a difference along the essential dimensions of the determinable. I think that Kagan was right, when making the analogy with volume, to notice that unpleasantness may vary along a dimension. However, we should notice that there might be multiple dimensions along which unpleasantness may vary.

Let us focus on the dimensions of being coloured. Three dimensions arguably constitute the property of being coloured: hue, saturation, and brightness. These three elements are necessary and sufficient for being coloured. Being coloured is equivalent to all the possible combinations along these three dimensions. Being coloured is the disjunction of the possible combinations of hue, saturation, and brightness. Accordingly, being red is the disjunction of the possible combinations of hue, saturation, and brightness, but within a more limited range than being coloured. For instance, a very precise shade of red, such as Coca-Cola red, is composed by a certain hue, saturation, and brightness, and the same applies to all the possible shades of red. The difference between two very specific shades of red, such as Coca-Cola red and Ferrari red, is that they have different hue, saturation, or brightness. This is what it means to be different non-additively, i.e., for properties to vary along the same essential dimensions.

I think that this accounts for what Johnson had in mind when he talked about a "special kind of difference". That is, the special difference between colour and shape is that they are different determinables, i.e., they have different essential dimensions. When we talk about different ways of being for various determinates of a common determinable, this means that the variation between these determinates is explained in terms of variation along *the same* dimensions. Contrast this to the different ways of being a young animal. This kind of difference is not explained in terms of variation along the same dimensions. The differences among these properties are explained in terms of variations along *different* essential dimensions: only the property of being a kitten entails being a cat, only the property of being a cub entails being a bear, only the property of being a lamb entails being a sheep, and so on.

If being a pain and being an itch are determinates of a common determinable, this means that these properties share the same essential dimensions. Being coloured is nothing over

and above the disjunction of all the possible ways of being red, blue, yellow, etc., i.e., being coloured is nothing over and above the disjunction of all the possible combinations of hue, saturation, and brightness. In a similar fashion, the phenomenal property of being unpleasant is nothing over and above the disjunction of all the different ways of being unpleasant, where these may include being a pain, being an itch, being a cramp, etc. These different ways of being unpleasant should be accounted for in terms of variations among the same essential dimensions.

Going back to Kagan's proposal, unpleasantness is not what volume is to sound. Rather, unpleasantness seems to be like sound in the sense that sound, as unpleasantness, may vary along multiple dimensions. Volume is one of the dimensions along which sound may vary. Similarly, intensity is one of the dimensions along which different ways of being unpleasant may vary. Sound may vary along dimensions other than volume, such as pitch, timbre, duration, etc. This also applies to different unpleasant experiences that might vary along dimensions other than intensity. However, which are the precise dimensions of unpleasantness? Before we can answer this, I think we need to consider a fundamental problem for the theory that we have developed so far.

4.3 Pains that are not unpleasant

I think that Crisp's idea of the determinable-determinate distinction as I have developed addresses the heterogeneity problem for unpleasant experiences. However, there is a critical problem for this account. This way of understanding pain in relation to unpleasantness entails that pain is necessarily unpleasant. However, it does not seem to be necessary that pain is unpleasant. Let me explain this. Determinate properties entail their determinables. For example, if being scarlet is a determinate of being red, being scarlet entails being red. Moreover, if being red is a determinate of being coloured, since scarlet entails being red, being scarlet also entails being coloured. If being a pain is a determinate of being unpleasant, then being a pain entails being unpleasant. This means that being a pain without being unpleasant is *impossible*, as it is impossible to be red without being coloured. However, there are cases that suggest that people might experience pains that are not unpleasant: pain asymbolia is often taken to be the more convincing example of this.

We have found a theory capable of dealing with the heterogeneity problem, while maintaining that unpleasantness is a phenomenal property; however, there might be actual cases that contradict the theory. This problem leads us to a fundamental question about the

nature of pain and unpleasantness: are there pains that are not unpleasant? I think we should benefit from the problem of understanding pain as a determinate of unpleasantness and see if there are actual cases that put into question that pain entails unpleasantness. I think that this quest needs to be taken seriously in order to have a detailed understanding of the nature of pain and unpleasantness. So, in the next chapters I will consider if there are in fact cases where people experience pains that are not unpleasant. If we find actual examples of people who have pain experiences that are not unpleasant, this shows that the theory about pain and unpleasantness that we have developed so far is wrong.

4.4 Conclusion

In this chapter I have explained two different theories that rely on the notion of dimension in order to account for the unpleasantness of pain. I have argued that the determinable-determinate distinction is useful to deal with the heterogeneity problem, while maintaining that unpleasantness is a phenomenal property. A key feature of this theory is that determinates of a common determinable vary non-additively, i.e., we can account for differences among different determinates of a common determinable by appealing to variations along their constitutive dimensions.

However, according to this theory, being a pain entails being unpleasant. It is necessary that if an experience is a pain, it is also unpleasant. In contrast, many philosophers consider that it is not only possible, but that there are actual examples of pain experiences that are not unpleasant. In the next chapter I take a detour from the theories of unpleasantness and I examine the candidates of pain experiences that are not unpleasant. In doing this, I address something that I take to be crucial to understanding the nature of pain and unpleasantness, i.e., the possibility of an experience being a pain without also being unpleasant.

CHAPTER 5: ARE PAINS ALWAYS UNPLEASANT?

5.0 Introduction

There is an important aspect to the relation between pain and unpleasantness. Given that pain is the paradigmatic unpleasant experience, we should ask: are pains always unpleasant? The question regarding whether pain is necessarily unpleasant has been important within recent philosophical discussion. In this chapter I will analyse various cases that suggest that pain might not be unpleasant. That is, these cases are possible examples of experiences that feel like a pain but that are not unpleasant. If these cases involve pain experiences that are not unpleasant, this shows that pain *is not necessarily* unpleasant. If being a pain does not entail being unpleasant, being a pain cannot be a determinate of unpleasantness, as I argued in Chapter Four. *Pain asymbolia* is the clearest example in which people experience pains that are not hedonic, i.e., pleasant or unpleasant. However, I will show that the evidence of the existence of pains that are not hedonic is not as conclusive as we might have thought.

I will organise the different pain disorders that might involve non-unpleasant pains into two categories. First, I will consider *pain insensitivity* cases. These are pain disorders that have been understood as cases where people are incapable of having pain experiences. There are different causal explanations for why these people do not experience pain, and different situations in which this may occur. I will describe various types of pain insensitivity cases and show that it is false that these people are *completely* incapable of experiencing pain. I will show that pain insensitivity cases *could* be interpreted as involving non-unpleasant pains, although they *shouldn't* be interpreted as such.

Second, I will consider cases of *pain indifference*. In these situations it seems as if people might become indifferent to their pain experience because their pains are not unpleasant. The standard analysis of pain indifference cases is to think that people stop reacting in the usual way to their pain experiences because such pains are not unpleasant anymore. However, I will show that this is not the only possible interpretation for these cases. There are two main alternative interpretations: i) to argue that people might be indifferent to their pain experiences, but that these experiences are still unpleasant, or ii) to argue that, in fact, what these patients feel are not really pain experiences. Pain asymbolia patients, also known as *asymbolics*, constitute the clearest candidate within this category.

I will conclude that even if there are cases suggesting the existence of pains that are not unpleasant, the evidence is not conclusive. The fact that pain is paradigmatically unpleasant suggests that all pains are always unpleasant *in fact*. In the next chapter, I will consider and explain the possibility of pain being pleasant rather than unpleasant.

5.1 Non-hedonic pains

Some philosophers have thought that unpleasantness is a *necessary* feature of pain: if it isn't unpleasant, it cannot be a pain. Some may think that "it is probable that if [pains] are all *in fact* unpleasant, then they are necessarily so..." (Pitcher, 1970, p. 491) In his 1970 paper, Pitcher considered various empirical cases that called into question the idea that pain is necessarily unpleasant. He considered cases of patients who had been through a lobotomy, who were fakirs, masochists, wounded soldiers, etc. He argued that these cases could be ultimately explained based on a pain model from the time - the Melzack and Wall's Gate Theory of pain (1965)[20] - and concluded that he could not find any example of a pain that wasn't *in fact* unpleasant. Even if people *seemed* to have pains that were not unpleasant, what was really happening is that the signal responsible for the pain experience was blocked, as the Gate Theory proposed, and this explained why in these cases there was no pain experience at all. This was meant to show that there are no actual cases of people experiencing pains that are not unpleasant. The fact that all pain experiences that we have are unpleasant is probably a consequence of pain being *necessarily* unpleasant, Pitcher thought.

However, not everyone agreed with him. Hall, for instance, argued that one "could have exactly the same kinds of sensation as you have when you are cut, burned, or bruised, and they not be unpleasant." (Hall, 1989, p. 643) Hall thinks it is at least *possible* that one could have all these normally unpleasant pain experiences, but without having the unpleasant aspect of them. These experiences would be pain-distinctive, i.e., they would be the kind of experiences that we are happy to call 'pain', hence not merely thermal experiences, or pressure experiences, etc., but without being unpleasant. The possibility of something like this implies that pain is not necessarily unpleasant: that it is possible that an experience could instantiate the phenomenal property of being a pain, without also instantiating the phenomenal property of being unpleasant. In the next sections I will consider various candidates for this kind of experience. If we find an *actual* example of an

[20] See reprinted version (Melzack & Wall, 1996).

experience that feels like a pain without also being unpleasant, this would confirm without a doubt that being a pain does not entail being unpleasant.

5.1.1 Congenital pain insensitivity

There are different conditions in which it seems as if people never experience pain. Let us start by focusing on *congenital pain insensitivity*, where people seem to be born without the capacity to experience pain. It is often thought that people with this condition *never* experience pain in their lives. They are insensitive in the sense that, whereas other people when stimulated with noxious stimuli would have a pain experience, people who are insensitive to pain would not have a pain experience at all. In this case, the insensitivity is congenital; people are born incapable of experiencing pain. They do not react in the usual way to stimuli that would normally cause pain and they claim that they do not experience pain in these situations.

> Thus, the defect must be *present from birth*, rather than acquired as a possible secondary manifestation of a disease process or traumatic injury; there must be a general *insensitivity to pain, i.e., an insensitivity to a variety of potentially noxious stimuli over the entire body, with no or slight involvement of the other sensory modalities*; and there must be no general mental or physical retardation. In short, persons with congenital insensitivity to pain must, strictly speaking, be "normal" in every respect other than this defect. (Sternbach, 1963, p. 253, my emphasis)

People with congenital pain insensitivity are rare and are more likely to suffer injuries, infections, and often die younger than the average population — which has made it difficult to fully understand the condition. It also shows the relevance of being able to experience pain as a means of protection and survival. However, it is worth noting that it is not obvious whether the protection role of pain is mainly or even fully dependent on its typical unpleasantness. Given that unpleasantness is taken to be motivational and normative, this suggests that the protective role of an unpleasant pain is importantly due to it being unpleasant. There are some cases of congenital pain insensitivity that indicate that it might be a hereditary condition; for example, some researchers propose that a possible cause for this condition is a genetic mutation of the SNC9A gene (Cox et al., 2006). However, regardless of the cause that explains congenital pain insensitivity, what is at stake for us now is what is happening with these people at the *experiential level*. Could it

be that what is actually happening is that congenital pain insensitivity patients experience pains that are not unpleasant? There are different possibilities that must be considered and explained.

It could be that congenital pain insensitivity patients are *completely* insensitive to pain, i.e., they have never undergone pain. If people with congenital pain insensitivity do not feel pain *at all*, this clearly isn't a case of pain that is not unpleasant, it is simply an example of people who do not have pain, either unpleasant or not. Why should we consider this condition then? I think this condition is relevant because it is not crystal clear that congenital pain insensitivity patients do not have any type of pain experiences in their whole lives. It seems that congenital pain insensitivity patients do experience pain *sometimes* in their lives. Sternbach (1963) refers to a few of these cases. For example, he talks of "a 34-year-old male, [who] had felt pain once when he had a smashed finger and once when kicked in the testes." (Sternbach, 1963, p. 254) Furthermore, the fact that congenital pain insensitivity individuals claim that they "don't feel pain" could actually mean different things.

This could mean that they are not conscious of any pain experience, while others would normally be conscious of feeling pain because something is happening to their bodies. Even if congenital pain insensitivity patients focused their attention and introspected while being burned or cut, there would not be any phenomenal pain experience that they could introspect. This seems partially correct, i.e., congenital pain insensitivity individuals do not seem to feel *anything* in some circumstances where others would feel pain. For example, Sternbach (1963) mentions the case of a woman who "was near death once from eclampsia at childbirth because she failed to recognise its symptoms (*she had no headache*)." (Sternbach, 1963, p. 260, my emphasis) Whereas other women would have consciously experienced a strong headache due to their physical condition, there was no introspectable headache for this woman, not even if she had focused her attention and tried to detect some pain experience of a headache.

However, on some other occasions, congenital pain insensitivity individuals do report feeling *something* when they receive stimuli that would normally cause pain. They seem to at least be able to detect and distinguish various tactile sensations. Researchers have noted that "detailed neurological examinations revealed that each [congenital pain insensitivity individual] could correctly perceive the sensations of touch, warm and cold temperature, proprioception, tickle and pressure, but not painful stimuli." (Cox et al., 2006, p. 895) This suggests that whereas many people would have an unpleasant burn

experience, for example, a congenital pain insensitivity individual would only experience some sort of thermoception, a certain warm feeling, which is not a pain.

Moreover, it is often reported that small children diagnosed with congenital pain insensitivity injure themselves by severely biting their fingers, lips, and tongue. However, with time they learn to avoid this damage, indicating that they are able to detect *the pressure* of their teeth against their skin. That is, these children have at least some form of tactile experience. For example, Juliao and Brotto (1955) reported a 3-year-old child who liked to play with fire, hit his head, and pull out his teeth. The only way of punishing him was by dousing him with a few drops of very cold water. This suggests that even if the child did not feel pain when pulling out his teeth, he had some form of tactile sensation when he was exposed to cold water.

However, I think that there is an important possibility that has been left unexamined. It could be that congenital pain insensitivity patients are really having a pain experience that is not unpleasant, but that they do not report it as being 'a pain' since pain is taken to be fundamentally unpleasant. Since these patients have this condition from birth, they do not develop the concept of pain, where 'pain' means, among other things, an experience that is unpleasant. Consider the child who pulls out his own teeth and plays with fire. It could be that he is having a phenomenal pain sensation that is not unpleasant, but given that it is not unpleasant, no one suspects that it is a pain experience. When congenital pain insensitivity patients feel something while being stimulated with noxious stimuli, but do not react with aversion and do not complain, this suggests that their experiences are not unpleasant, but it does not entail that what they feel is not a pain-like experience.

I think that there are two key points to conclude regarding cases of congenital pain insensitivity. First, contrary to what we might think, some people diagnosed with this condition do claim to experience pain in rare occasions in their lives, and these pains seem to be unpleasant — even if it also true that on some occasions congenital pain insensitivity patients do not seem to have any conscious experience, while others would experience an unpleasant pain. Second, it *could* be that congenital pain insensitivity patients experience pains that are not unpleasant, but that they do not report them as such because they learned that the word 'pain' is only used for unpleasant experiences. However, this possibility is not useful if we are looking for *actual* examples that confirm that people experience pains that are not unpleasant. Even if possible, there is no evidence suggesting that congenital pain insensitivity individuals have pain experiences that are not unpleasant. Even if congenital pain insensitivity cases do not provide us with evidence that people do

experience pains that are not unpleasant, there are other cases of pain insensitivity that are worth looking at.

5.1.2 Acquired pain insensitivity

There are other cases that suggest that people have lost their ability to experience pain, as against being born with that inability. They do not have a pain experience in the presence of noxious stimuli, or when their bodies are injured, as normal subjects would. These are the cases of *acquired pain insensitivity*. In these cases, it is due to different types of trauma or disease that people seem to stop being able to experience pain. However, as with congenital pain insensitivity, could it be that what is really happening with these cases is that people experience pains that are not unpleasant? In these cases people are normally capable of identifying pain experiences, so I think we should search in these cases for evidence of pains that are not unpleasant.

One interesting example of acquired pain insensitivity is the injured soldiers who do not seem to experience pain right after being in the battlefield. Beecher (1956) brought attention to these cases through his work on injured soldiers in the Anzio Beachhead during World War II. Beecher was surprised by the difference between these soldiers and the general public; he wanted to explain why these injured soldiers did not complain about being in pain, even if they were severely injured, while civilians would normally complain much more about being in pain, even if they had less severe wounds. Beecher reasoned that the difference could be explained in virtue of a psychological factor: soldiers did not complain about their pain because they had some sort of analgesia due to the fact that the wounds meant something *positive* for them. Since the wounds meant something positive for these soldiers, the soldiers stopped feeling pain.

In Beecher's study the soldiers were asked the following question: as you lie there, are you having any pain? And most soldiers answered 'no'. From this, Beecher concluded that these soldiers said 'no' because they did not feel any pain in relation to their war wounds. That is to say, this meant that soldiers were not aware of any pain experience in relation to their injury, in a similar way to which congenital pain insensitivity patients seemed to be unable to have any conscious pain experience when injured. The soldiers were not conscious of any unpleasant pain in relation to their recent bodily damage from combat, even after careful introspection. Indeed, it does seem as if sometimes these soldiers did not have any conscious pain experience in relation to their war wounds.

According to Beecher, these wounds meant something positive for the soldiers - that they were now out of the battlefield and could return home - which explained their lack of pain.

Not everyone agrees with Beecher's explanation of why the soldiers didn't feel pain. Klein points out that Beecher's explanation is implausible given that "painless injury is also common even among people who view their injuries as *entirely negative*." (Klein, 2015b, p. 28, my emphasis) For example, a study on Israeli soldiers with traumatic amputations after the Yom Kippur War showed that wounded men reported that their initial injury was also painless even if they saw the consequences of their injury as *negative*. When talking about the injured Israeli soldiers, Wall (1999, p. 8) says that "[t]here was never a hint that anyone adopted the Darwinian approach that being wounded increased their chance of survival. No soldier reported a fleeting sense of relief that they have escaped alive from the killing fields." That is to say, one does not have to interpret one's wounds as positive in order to have some sort of analgesia regarding those injuries. On the contrary, one might experience analgesia even if the wounds mean something rather negative.

Regardless of what the right explanation might be for why these soldiers were not feeling pain, we should verify if it is true that these soldiers really stopped experiencing any type of pain and, more precisely, any pain in relation to their battle wounds. In fact, the soldiers do not seem to have a general analgesia for any type of pain experience and, more importantly, it seems that they actually might experience pain in relation to their war wounds. As Beecher acknowledges, "these severely wounded men did not have a general block of the pain experience, for they complained in a normal manner at rough handling of their wounds, or at inept venipunctures [i.e., when they took a blood sample from the soldiers]." (Beecher, 1956, p. 1610) The soldiers' complaints suggest that they were not *completely* incapable of experiencing pain, since they complain about badly taken blood samples. Moreover, since they complained "at rough handling of their wounds", this may imply that they did experience some type of unpleasant pain in relation to their battle injuries. But if the soldiers seemed to be capable of having unpleasant pains, why did they answer that they didn't feel pain?

Here's another possibility for what was happening. These soldiers still felt an unpleasant pain, but their pain was so mild that they were not bothered by it, they didn't ask for treatment, or complain about it.[21] When someone asked them if they felt pain and they

[21] Bain (2013) suggests something along these lines. The soldiers feel unpleasant pains, but less unpleasant than a civilian would with the same injury.

said ‘no’, what this meant was ‘no, not enough to be complaining about it’. If this interpretation were correct, then soldiers’ pain was still unpleasant, even if it was not very intense. This interpretation, however, might be implausible since it would be inconsistent with the fact that soldiers do complain about other things such as the badly performed venipunctures. If these soldiers complained about the unpleasant experience caused by the badly performed extraction of blood, why wouldn’t they also complain about their war wounds if they caused them some kind of mild unpleasant pain experience? One possible explanation is that not complaining about battle wounds is more honourable that not complaining about a badly performed blood sample, or about the rough handling of their wounds.

In any case, this is not what we are concerned with at the moment. The possibility we are interested in is whether soldiers answered that they did not feel pain, but what was really happening was that they had a pain experience that was not unpleasant, so they thought they did not need treatment for it. This, however, seems rather unlikely. It is possible, but not really plausible. There is no behavioural or testimonial evidence of the soldiers reporting that they do feel pain, but that it is not unpleasant at all. Even if it is *possible* that people with acquired pain insensitivity had pains that were not unpleasant, there is no evidence suggesting that this possibility is *in fact* correct.

In conclusion, it does not seem like acquired pain insensitivity provides us with an example of a pain experience that is not unpleasant. It seems like people might enter into some state of complete analgesia in relation to major wounds. There might also be other forms of analgesia during the heat of the battle, where people might become incapable of experiencing any pain. However, it is hard to run a questionnaire about what exactly people feel in these conditions, for obvious reasons. In any case, beyond the explanation of why soldiers would have analgesia for some pains but not for others, we have not encountered yet a conclusive case of a pain sensory experience that is not unpleasant. Let us now look into cases where people do claim to experience pain, but do not show the normal reaction to it.

5.1.3 Pain indifference due to surgery

Some people do report experiencing pain without showing the typical reaction to it: they seem to be indifferent to their pain. They are indifferent in the sense that they stop having the typical reactions of avoidance to their pain experience, and stop showing the typical

behaviour in relation to pain, such as trying to eliminate it, even if they claim that they experience pain. This happens, for example, in cases of *lobotomy*. Sometimes people used to be given lobotomies when they suffered from a very intense chronic pain. In this procedure, also called *prefrontal leucotomy*, the nerve fibres of the frontal part were disconnected from the rest of the brain. Lobotomy cases have made philosophers consider whether these patients have pains that are not unpleasant (e.g. Dennett, 1978; Pitcher, 1970). Lobotomy patients recognised their experience as 'the same' before and after the surgery, but for some reason they stopped showing distress and claimed that they were not bothered about the pain anymore after the procedure. In other words, it seems as if they became *indifferent* to their pain. This indifference is often explained as the consequence of pain not being unpleasant anymore.

For example, Freeman and Watts (1946, p. 995) reported a woman who, after the operation, said that the pain was "exactly the same as it was before"; then, when they asked her about the fact that she was not complaining about the pain anymore, she answered that she "couldn't do anything about it, so it doesn't do any good to complain". The same authors reported another case where, a year after the operation, "[t]he patient stated that the pain was just as it was before the operation. And yet she does not talk about it, she is rather carefree, and her arm is a nuisance, rather than a constant reminder of her permanent invalidism." (Watts & Freeman, 1948, p. 717)[22] How should we interpret the fact that these lobotomy patients say that the pain remains *the same* and yet they react so differently after the surgery? We could argue that something in their pain experience changed; even if they say that the pain is "exactly the same as it was before" and that it "was just as it was before the operation", what they mean is that the pain remained the same *qua* its sensory phenomenology, yet that sensory pain experience lost its unpleasantness. This explains why "it doesn't do any good to complain" and why the patient becomes "rather carefree".

In contrast to the previous cases of congenital pain insensitivity, I think we can rule out the possibility that lobotomy patients do not feel any conscious pain experience. Their verbal reports and some of the patients' behaviour strongly suggest that they do feel *something,* and that what they feel is a pain experience. They were able to experience and distinguish pain experiences before the operation and they claim that they can still feel a pain after the surgery. Moreover, there is no evidence suggesting that they became incapable of using the term adequately, it is not as if they were completely incapable of

[22] As cited in Trigg's (1970, p. 131).

reporting that they have conscious phenomenal experiences. If they can report having other sensory experiences then there is no reason to think that they cannot report their pain. However, even if we do accept that these lobotomised patients still feel something, that they can feel pain, there are other ways of understanding what it is that they feel.

First, it could be that *all* their pains stop being unpleasant after the surgery. Could this interpretation be correct? I don't think so: lobotomy patients seemed to still complain about other pains. Even if there is not much evidence about the patients' reaction to sources of pain that were not related to the ones for which they were lobotomised, "[i]t is certain that the sensory component of pain is still present because these patients may complain vociferously about pinprick and mild burn." (Melzack, Wall, & Melzack, 1982, p. 169) Moreover, "they respond or sometimes *over-respond* to a normally painful stimulus." (Melzack et al., 1982, p. 296, my emphasis) Second, it could be that only their chronic pains, the ones for which they were lobotomised, stop being unpleasant. Could this interpretation be correct?

Lobotomy patients still sometimes complained when suffering from episodes of their chronic pain. For example, Freeman and Watts (1946, p. 955) report the case of a woman ten years after the procedure that "when she is asked about her sensations she puts on a long face and tells how terribly she feels, how much her back hurts, how she can hardly walk; and yet she never complains of these sensations to members of the family". The fact that this patient does not complain *to her family*, does not show that she does not complain at all. After all, she says that she feels terrible. In fact, lobotomised patients do continue to complain about their chronic pain, even if their attitude changes about it. Given that patients seem to continue suffering, how can we explain their behavioural changes?

I think that it is more likely that their pains are still unpleasant. Trigg (1970) offers a possible answer to explain why there might be behavioural changes in lobotomised patients about their chronic pain. He points out that lobotomised patients seem to have lost the capacity *to fear and to be anxious* about their pain. Given that the anxiety towards experiencing strong pain episodes seems to have been the main source of suffering for some of these patients, this would explain why they acquired indifference, i.e., why they complain so much less about their condition after the procedure. If lobotomy patients lost the capacity to fear and be stressed about their pain due to the lobotomy, they can now stand their still unpleasant pain and be somewhat indifferent about it, which does not imply that their pains are not unpleasant at all. Moreover, in line with Trigg's explanation, Freeman and Watts claim that:

> Prefrontal lobotomy relieves anxiety and emotional tension in rather specific fashion. It diminishes concern over consequences. It eliminates the fear of pain. Pain may be present; but, when it no longer arouses a mental picture of future disability and all that this may mean in terms of disaster to the person and his family, *the experience can be borne with equanimity*. Consequently, prefrontal lobotomy lends itself well to the relief of conditions in which the affective component of a painful disorder is equally important with the local condition… Psychosurgery alters the subject's reaction to pain *without materially changing his ability to feel pain*. Pain may be present, but when divorced from its implications —insecurity, disability, guilt, death— it then becomes bearable and may be accepted with fortitude. (Freeman & Watts, 1946, p. 954–955, my emphasis)

This is a clear statement that confirms Trigg's proposal. Given that the surgery does not change their ability to feel pain, why should we think that lobotomy patients' pain has lost its unpleasantness? The main constant feature among lobotomised patients is that they lose some emotional response, such as stress and anxiety, to their chronic pain, which is what explains their change in behaviour, what accounts for their indifference. When Freeman and Watts talk about a lost "affective component", this does not necessarily imply that pain has lost its felt unpleasantness, instead, it may mean that people have lost the affective emotions in relation to their still unpleasant pains.

I think that Trigg's proposal is convincing given that, in fact, lobotomised patients claim that their pains, when felt, remain the same. They do not claim that they feel the same type of pain but without their normal unpleasantness. People who reported on lobotomy cases also emphasise that a huge part of patient suffering was emotional; researchers describe that the clearest change after the surgery is in people's attitude to their pain, not in the pain experience itself. It seems to me that this is a more consistent and intuitive interpretation than saying that lobotomised patients lost the felt unpleasantness of their pains.

Moreover, other similar cases can be explained in the same fashion. This explanation can also be applied to cases where people stop having averse reactions to their pain experiences. For example, "patients who have been treated with morphine because of severe post-operative discomfort or extreme pain from cancer frequently tell their doctors, 'It's a funny thing. The pain is still there, but it doesn't bother me.'" (Snyder, 1996, p. 44) There is evidence suggesting that morphine relieves pain by relieving anxiety (Hill, Belleville, & Wikler, 1955). That is to say, in this case we could also make sense of the

change in the response to pain by attributing a change in the levels of anxiety, which is a possible effect of morphine use.[23] Once more, the change is not in the phenomenal qualities of the pain. Rather, the change in anxiety explains why we can have such a different reaction to a still unpleasant pain.

Another candidate for this kind of indifference can be found in patients who are hypnotised to treat their pain during surgery or other painful procedures. These hypnotised patients still describe the experience as a persisting pain, but now experience that pain from a distance and with no suffering or concern (Barber & Bejenke, 1996, p. 90–91). I take it that the felt distance and the lack of concern are rather explained in terms of changes in people's attitude towards their felt experience, rather than a change in the phenomenology of the experience itself. This is not to say that it is impossible that hypnosis only takes away the unpleasant aspect from the sensory pain, but it is a less likely explanation of what is happening. I think it is more likely that hypnosis changes our attitude towards an unpleasant pain experience; we are able to change certain mental states such as beliefs and desires through hypnosis, and as a result of this we embrace our unpleasant pain. If we still have a conscious experience of pain, by thinking that it is fine to have that experience, or by managing to want it, we can obtain a different and much less aversive reaction to an unpleasant pain, especially if it is mildly unpleasant. This seems to be a more likely interpretation of what hypnosis is doing than thinking that hypnosis changes the phenomenology of an unpleasant pain by only vanishing its unpleasantness away.[24]

All in all, I think that the most plausible interpretation for lobotomy cases is that if these people still have conscious experiences of pain, then these experiences remain unpleasant. What explains the change in people's behaviour, their indifference, does not come from a change in the pain experience, but from a change in their emotions about those still unpleasant pains. I think that this is highly consistent with the reports from lobotomised patients. It is also a plausible explanation of the effects of morphine and hypnosis, and it is in line with pain being the paradigmatic example of an unpleasant experience. In the face of the evidence, I do not think there are enough reasons to drop the very strong intuition that pains are actually unpleasant. That said, I think it is time now to address the strongest candidate of a condition where people might have pains that are not unpleasant.

---

[23] It is worth pointing out that Klein (2015a) defends that morphine cases and pain asymbolia should be understood similarly. I explain his view on pain asymbolia in the upcoming section of this chapter.

[24] For more on the effects of hypnosis on the unpleasantness of pain a see Rainville et al. (1999).

5.1.4 Pain indifference due to lesion: pain asymbolia

Pain asymbolia is a neurological condition that has drawn a lot of attention in the philosophical discussion about pain (see Bain, 2013; Corns, 2014; Grahek, 2007; Gray, 2014; Klein, 2015a; Vignemont, 2015). This condition, some philosophers think, provides strong evidence of the existence of pains that are not unpleasant. In other words, there might be mental states that instantiate the phenomenal property of being a pain but without instantiating any phenomenal property of unpleasantness. Asymbolics do not react in the usual way to harmful stimuli that would normally cause pain, yet they claim that they feel pain. Moreover, they say their pains do not hurt and sometimes start smiling or laughing when such noxious stimuli is applied to them. Schilder and Stengel (1928) provide the case of a woman diagnosed with pain asymbolia.

> [She] displays a striking behavior in the presence of pain. She reacts either not at all or insufficiently to being pricked, struck with hard objects, and pinched... Pricked on the right palm, the patient smiles joyfully, winces a little, and then says, "Oh, pain, that hurts". She laughs, and reaches the hand further toward the investigator…" (Schilder & Stengel, 1928, p. 147)[25]

This rare condition is explained as the result of a *lesion* in the posterior insula, typically caused by strokes and brain tumours in adulthood (Berthier, Starkstein, & Leiguarda, 1988). Grahek (2007) brought philosophical attention to these cases; he thinks that pain asymbolia proves the existence of pains that are not unpleasant. According to Grahek, normal pain experiences have many components. Among these, there is a *sensory* component, which is associated with the detection of stimuli harmful for the body, and an *affective/motivational* component, which constitutes the unpleasant and motivational aspect of pain. That is, an unpleasant pain is composed of: i) the sensory pain aspect and ii) an unpleasant aspect. Grahek thinks that people with pain asymbolia have the former aspect but lack the latter, i.e., they do have a pain experience, but their pain is not unpleasant and, therefore, they are not motivated to avoid the stimuli responsible for their pain. In other words, asymbolics have a mental state that only instantiates the sensory aspect of pain, and given that it is in virtue of being unpleasant that a normally unpleasant pain motivates action, the lack of unpleasantness explains the lack of avoidance behaviour to harmful stimuli.

[25] From Klein (2015b, p. 142).

Not everyone agrees with Grahek. Klein (2015a) argues for a different interpretation of pain asymbolia. Klein thinks that asymbolics' lack of motivation in the presence of pain experience is due to losing "a general capacity to care about their bodily integrity" (Klein, 2015a, p. 493). That is to say, the lesion caused after a stroke or a tumour is responsible for a general lack of care about one's own bodily integrity. Asymbolics do not care about their body so they are not motivated to act in relation to their pains, even if they are still unpleasant; "patients are indifferent to pain not because the pain has changed, but because they no longer appreciate it as a command worth following." (Klein, 2015a, p. 512) Another way of putting this is that care is a necessary condition for motivation, and since asymbolics lack the relevant care, their unpleasant pains are not motivational. If asymbolics have lost the care for their body integrity, the idea goes, they have also lost the care for the unpleasantness of their pains.

Klein's explanation of pain asymbolia is similar to Trigg's regarding lobotomy. That is to say, both think that pains are felt in the same way as usual. The explanation in Trigg's and Klein's proposals consists in a change in the patients who have pain experiences that are still unpleasant, not in the experience itself. According to Trigg, the lack of anxiety and emotional response explains the indifference of lobotomised patients to their unpleasant pains; Klein thinks that the lack of bodily care accounts for the indifference in asymbolics to their own unpleasant pains. This general lack of bodily care, Klein thinks, is confirmed by asymbolics' behaviour more generally. This behaviour is sometimes neglected in the explanations of pain asymbolia. For example, asymbolics fail to show a normal response when matches are struck close to their face and eyes, and when they are presented with loud noises or strong flashes. Klein argues that his explanation is better that Grahek's because the lack of bodily care also explains asymbolics' general lack of motivation to avoid many threats and dangerous situations. According to Klein, asymbolics' general lack of care about the integrity of their bodies also explains that having matches struck close to their face is not enough to motivate action.

However, not everyone agrees with Klein. Namely, we could argue that precisely because asymbolics lack that care, their pains stop being unpleasant (Bain, 2013). Bain thinks that pain is not unpleasant for asymbolics precisely *because* they have lost their capacity to care about their body integrity, which also explains more generally their strange behaviour in the presence of danger. Let us recall that Bain defends a form of mental content theories, BAP, as we explained in Chapter One. He defends a form of evaluativism, the idea that pain's unpleasantness consists in representing a bodily disturbance as being bad; a pain is unpleasant insofar as the bodily disturbance is represented as being bad.

According to Bain, Klein is right to think that asymbolics have lost their capacity to care about their body integrity, which thus explains why they don't care about bodily damage either and why they do not react normally in the presence of dangerous stimuli. However, this lack of care also implies that asymbolics stop representing their bodily disturbances as being bad. That is to say, pain, for asymbolics, *stopped being unpleasant*.

De Vignemont (2015) offers an interesting turn. If she is right, both Klein and Bain must be wrong. She argues that the lack of bodily care does not imply that pain stops being unpleasant for asymbolics. De Vignemont discusses the cases of *somatoparaphrenia*: people who experience pain in an 'alien' part of their body, i.e., a part that they do not identify as their own. These patients, she argues, do not care about the alien part of their body or about the pain felt in the alien part of their body. In short, somatoparaphrenia patients do not care about the integrity of their alien body parts. However, they still feel pain and are motivated to avoid it, which implies that they still have unpleasant pains and are thus motivated in relation to this experience. For example, a somatoparaphrenia patient may identify her right hand as not being really hers; however, even if this patient says that such a hand is not really hers and that the felt unpleasant pain isn't hers either, she nonetheless still acts in order to protect the hand from the stimuli that causes the unpleasant pain.

If this is correct, it challenges Bain and Klein's idea that the lack of bodily care explains asymbolics' strange behaviour in relation to pain: somatoparaphrenia cases show that one could experience an unpleasant pain and react normally to it, even if one does not care about one's own body. In others words, the somatoparaphrenia cases that de Vignemont points out show that Bain is wrong, i.e., care is not necessary for having unpleasant pain experiences, one might not care and still have an unpleasant pain. This goes against Bain since he thinks that if we do not care about our own body, then the pain is not unpleasant. Somatoparaphenia patients show exactly the opposite: one could have an unpleasant pain without caring about one's own body. Somatoparaphrenia also shows that Klein is wrong, i.e., care about our body integrity is not necessary for motivation in relation to an unpleasant pain. Klein thinks that if one does not care, then one won't be motivated, however, somatoparaphrenia cases show the opposite. Somatoparaphrenia patients do not care about their own body and yet they are still motivated to avoid their pains.[26]

[26] See Klein (2017) for a recent response to de Vignemont's critique regarding motivation. Even if Klein's response were correct, this does not entail the existence of non-unpleasant pains.

So where are we? How can we account for asymbolics' odd behaviour in the presence of noxious stimuli, together with the fact that they claim to feel pain? If de Vignemont's mention of somatoparaphrenia undermines Klein and Bain's interpretations of pain asymbolia, should we conclude, as Grahek did, that asymbolics experience pain without unpleasantness? Is there another alternative explanation for pain asymbolia? Yes, it could be that asymbolics do not actually experience pain *at all*. Gray (2014) proposes this possibility based on an *intensive theory of pain*. He defends that pain is not a sensory modality like vision, hearing, touch, etc., that provides us with information from physical stimuli. Instead, pain "has the role of warning us of excessively intense physical stimuli." (Gray, 2014, p. 95) We experience pain when we receive stimuli through other sense modalities, and when we represent that the stimuli are so intense that it would normally cause harm, then we feel pain.

For example, when we represent an intense change in temperature or pressure, we experience pain. More precisely, a pain is the representation of a change in temperature or pressure that is so intense that it might result in bodily damage. This could be understood as what constitutes the phenomenal property of being a pain. Based on this account, Gray (2014) thinks about asymbolics that:

> The threshold at which stimuli cause pain is raised in asymbolics such that stimuli that previously exceeded the threshold and triggered the experience of pain no longer do so. Asymbolics still think of their experiences in terms of pain because they remain very intense. (Gray, 2014, p. 95)

In other words, we could understand pain asymbolia as an example of *acquired insensitivity*. That is, if we accept Gray's interpretation of pain asymbolia, we could explain that asymbolics were able to experience pain, but due to a lesion after a stroke they become incapable of experiencing pain. Their threshold for detecting that a stimulus is dangerous is so high that they do not experience pain anymore. They become unable to represent when stimuli detected in other sensory modalities are so intense that they may lead to injury. Moreover, if we accepted Gray's account of pain asymbolia, we could explain that this is also coherent with asymbolics' general strange behaviour: they do not react to other dangerous stimuli because they have lost a general capacity to represent danger. One of the results of this general loss of the ability to represent damage is that they cannot feel pain, since having a pain consists in representing when stimuli from other sensory modalities would cause injury, viz., when the stimuli becomes dangerous. In

short, if we accept Gray's proposal, we could argue that pain asymbolia is actually a case of acquired pain insensitivity.

However, we could still wonder: why do asymbolics claim to feel pain while other pain insensitives do not? People who were born with congenital pain insensitivity and the soldiers coming from the battlefield have said that, at least generally, they did not feel pain. In contrast, asymbolics clearly say that they do. How can we explain such difference? I think we can answer that asymbolics *confuse* their intense experiences of other sensory modalities with pain. In contrast to other cases like congenital pain insensitivity, asymbolics could normally experience pain before they suffered their lesion. They were used to having intense experiences that resulted in unpleasant pains, and now they still have experiences as intense as the ones that resulted in pain, but without the pain that used to accompany these intense experiences from other sensory modalities; they continue to call intense experience like this 'pain'.

Let me try to be clearer. The idea is that asymbolics confuse: i) intense experiences that are not unpleasant coming from another sensory modality such as pressure or thermoception with ii) having a pain that is not unpleasant. They call the intense experience of another sensory modality 'a pain that is not unpleasant'. They get confused because such intense experience of touch, say, used to be accompanied by an unpleasant pain. For instance, when an asymbolic was burned before the lesion, she would feel an intense thermal experience, a sensation of warmth, and eventually a pain. However now, after the lesion, she only continues to feel the intense sensation of warmth. This translates into her saying that she still has a pain that is not unpleasant. If this looks like an eccentric explanation, pain asymbolia is a very uncommon condition, so it shouldn't be too shocking to have an odd solution for a strange phenomenon.

There is a remaining question to be asked: how do we explain the difference between pain asymbolia and cases of acquired pain insensitivity such as in the wounded soldiers? Why don't these soldiers also say that they feel pain? I think we can give an explanation for this. Let us recall that the wounded soldiers only stopped feeling pain in relation to their major wounds. However, for obvious reason, the doctors and nurses did not apply intense noxious stimuli to the areas where these soldiers were significantly injured to investigate if they still felt pain. If they did, maybe the soldiers would have given similar reports to the asymbolics' and say that they felt pain in those areas, but that it was not unpleasant. However, in the lack of evidence, I think that it is a plausible explanation to argue that pain asymbolics do not feel pain.

I think that in order to decide how to best interpret pain asymbolia, we should recapitulate what has been discussed so far. We have seen four interpretations of pain asymbolia: i) Grahek's, who thinks that asymbolics have pains that are not unpleasant, which explains their odd behaviour; ii) Klein's, who thinks that they still have pains that are unpleasant, but that they lost their capacity to care about their body integrity which explains their odd behaviour; iii) Bain's, who thinks that asymbolics' pains are not unpleasant, but they are not unpleasant because they have lost their capacity to care; and iv) Gray's, who thinks that asymbolics have lost their capacity to feel pain, which could be explained, I think, as a general lost capacity to represent danger. Klein's and Bain's approaches were disqualified as possible explanations by de Vignemont's cases of somatoparaphrenia, that show that care is not necessary for motivation and it is not necessary for pain being unpleasant either. So we have to decide between Grahek's view, and what I take to be a more developed version of Gray's approach.

I think that we should opt for the latter for a simple yet strong reason: pain is the paradigmatic example of an unpleasant experience. That is to say, there is a very strong intuition that pain might entail, or at least usually involve, unpleasantness. If *X* is paradigmatically an example of *Y*, we should have very strong evidence to make us change our mind. However, I don't think we have very strong evidence to accept the actual existence of pains that are not unpleasant. Let me offer an analogy. I take it that *rape* is a paradigmatic example of *something wrong*, something that should not be done. That is, it is very hard to find a scenario where rape can be conceived as something that is not wrong. Maybe one could say that, in a very utilitarian way, if we could save humanity at the cost of one rape, then that rape would not be that wrong. However, we could argue that, all else being equal, i.e., beyond some ultimate calculation of costs and benefits, rape is in itself wrong. The main idea is that if we have the intuition that rape is a paradigmatic example of something wrong, we have to be confronted with a very strong counterexample to be willing to drop our intuition. In analogy, we would have to be confronted with quite indubitable evidence of a pain experience that is not unpleasant to accept that this actually occurs. After careful examination of possible cases of pain that might not be unpleasant, I do not think that the evidence is strong enough to renounce the intuition of pain being in fact always unpleasant.

Finally, even in pain asymbolia, which is meant to be the clearest example of a pain experience that is not hedonic, Schilder and Stengel's patient says "Oh pain, that *hurts*". That might imply that, whatever she is feeling, it is unpleasant, and thus shows that it cannot be a pain that is not unpleasant. So, whether or not she is feeling a pain, as long as

her experience is unpleasant, it is not a case that puts into question the paradigmatically unpleasant nature of pain. Pain asymbolia seems to be the clearest example of a pain that is not hedonic, but the evidence is not conclusive about asymbolics feeling pains that are not unpleasant.[27]

5.2 Conclusion

In this chapter I have analysed the possibility of experiences instantiating the property of being a pain without also being unpleasant. After careful consideration, I think that there is not strong enough evidence to renounce the intuition that pains are always, in fact, unpleasant. This is supported by the general idea that pain is the paradigmatic example of an unpleasant experience. This analysis allows us to have a better understanding of different situations where people experience pain and the relation between pain and unpleasantness. There is one last important candidate of a pain experience that is not unpleasant that should be considered. So far we have examined the possibility of a pain experience not being hedonic, i.e., being neither pleasant nor unpleasant. In the upcoming chapter I will focus on masochism and the possibility of having pains that are pleasant rather than unpleasant.

[27] There is an interesting possibility of unpleasant experiences without having the normally associated pain aspect. Ploner, Freund, and Schnitzler (1999) present a case like this. I will go into more detail about this possibility in Chapter Seven.

CHAPTER 6: MASOCHISM

6.0 Introduction

Masochism is an interesting case for understanding the nature of pain and unpleasantness. It is interesting in the sense that it seems to entail two inconsistent things: i) pain is unpleasant and we should *avoid* it because of such unpleasantness, and yet ii) masochists are *attracted* to pain experiences, they pursue these experiences even if they are unpleasant. How can we explain that masochists seek out pain experiences if, as we established in the previous chapter, it seems that all actual instances of pain are unpleasant? Are there instances of pains that are only pleasant? If there are examples of pains that are only pleasant, this would explain why masochists are attracted to these pain experiences and show that there are non-unpleasant pains.

I will delineate different scenarios where people *seek out* pain experiences and make sense of them. I will refer to all these as masochistic cases. In doing this, I will show that the term 'masochism' has been used to refer to many different circumstances. My main aim is to describe the different scenarios where people seek out pain experiences, and explain the reasons why they might do so. By analysing these cases, we will understand the reasons why we might intentionally pursue pain, and confirm whether there are instances where people seem to experience pains that are only pleasant, i.e., that people experience pains that are non-unpleasant.

First, I will discuss *means-ends masochism*. This is the most common explanation of what is occurring during scenarios where people seek out pain. Broadly, people might pursue pain *even if* it is unpleasant because it is the means for something else that is what is really being pursued. I will argue that these are cases where the pain is pursued *because* it is unpleasant. I think that there are two clear examples of this in the literature about masochism: i) when a pain is pursued as a means for the feeling of submission, and ii) when a pain is pursued as a means for punishment.

Second, I will consider *side effect pains*. These cases might seem very similar to the previous category. However, in these situations people have pain experiences without the pain having an instrumental role. Going to the dentist is a typical example where this occurs, i.e., we might have to *put up* with an unpleasant pain in order to have better dental health, but it is not in virtue of feeling this pain that our health improves. The pain is a side effect, it does not play an instrumental role in itself; it is rather collateral damage. I do

not think that these are cases of masochism because the pain experience is not sought out, i.e., the pain is not intentionally pursued. I think these cases should be pointed out so they are not confused with means-end masochism. For example, some cases that are considered as sexual masochism, I think, fall under this category.

Third, I will give an account of *end-in-itself masochism*. In these cases pain is pursued as an end, which might suggest that it is not unpleasant. I will divide this category into two sub-categories. First, I will explain *context masochism*. In these cases a pain might be pursued as an end in the sense that the pain is an essential part of a whole. It is the whole that is sought out, strictly speaking; however, since a pain partially constitutes the whole, then the pain is also pursued as an end. A good example of something like this is, I think, when we want our food to be spicy. The second type of end-in-itself masochism is *masochistic pleasures*. These cases can be understood as people having token experiences that are pleasant and unpleasant simultaneously (Klein, 2014). According to Klein, the unpleasantness of a pain might be pleasant under certain circumstances.

Different authors have proposed various ways of understanding what masochism is, but I think that what they were really doing was to describe different types of masochism, different versions of a common phenomenon. The common feature in all these types of masochism is that people seek out the feeling of pain. However, the reasons why they seek out these experiences may differ. Moreover, it does not seem like people seek out pain experiences that are *only* pleasant. There is no conclusive evidence of non-unpleasant pains in the literature on masochism.

6.1 Means-end masochism

The most common attempt in the philosophical literature to explain masochism is to argue that masochists do not really want to have a pain for itself. Rather, what is really happening during masochism is that someone may pursue pain "as a means to some end... backed up by dark reasons like guilt or sexual masochism." (Nagel, 1986, pp. 156–157) In this first section I will show different ways in which pains may have instrumental value, which explain why we engage in such experiences *because* they are unpleasant.

The term ‘masochism’ was first introduced in medicine by Kraft-Ebing (1892) as a psychiatric condition, together with other terms such as ‘homosexuality’ and ‘fetishism’. The word was inspired by the name of the writer Leopold Sacher-Masoch. Severin, the

main character in Sacher-Masoch's book *Venus in Furs* (1870),[28] could be characterised as what Kraft-Ebing called a masochist. The fact that Kraft-Ebing was creating a classification of *abnormal* sexual behaviour is responsible, I suspect, for associating masochism with some form of pathological sexual behaviour. According to Kraft-Ebing, the key feature of masochism is not that pains are a means for sexual gratification, as masochism is often understood. Instead, pains were a means for feeling submissive.

According to Kraft-Ebing, masochism is "a peculiar perversion of the psychical vita sexualis, in which the individual affected, in sexual feeling and thought, is controlled by the idea of being completely and unconditionally subject to the will of a person of the opposite sex; of being treated by this person as by a master —humiliated and abused." (Krafft-Ebing, 1892, p. 89) In the original understanding of masochism, a masochist is not someone who seeks out pain for its own sake, pains is rather sought out as a means for something else. I think that the key feature in these cases is that it does not seem like pain experiences are not unpleasant. On the contrary, these cases suggest that the pains involved must be unpleasant. It is important that the pains are unpleasant because this helps to achieve the feeling of submission. When masters inflict pain on masochists, masochists would not feel that that their masters are really controlling them, that they are being humiliated and abused, if their pains were not unpleasant. For masochists to feel submissive, it is important that the pains inflicted on them are unpleasant.

Kraft-Ebing distinguished masochism from *sexual bondage*; the latter is, in fact, much closer to what the current medical definitions take masochism to be, and what we often have in mind when we think of masochism. Sexual bondage was understood as the use of pain as a means for some form of sexual satisfaction. According to Kraft-Ebing, in sexual bondage people use pain, or other unpleasant experiences, as a means for sexual pleasure. In his words:

> Sexual bondage is not a perversion and not pathological; the elements from which it arises - love and weakness of will - are not perverse… In masochism, which is decidedly abnormal and a perversion, this is all very different… I repeat that the decisive points, in the differentiation of simple passive flagellation from flagellation dependent upon masochistic desire, are that, in the former, the act is a means to make coitus, or at least ejaculation, possible;

[28] See reprinted version (Sacher-Masoch, 2004).

> and that, in the latter, it is a means of gratification of masochistic desires. (Krafft-Ebing, 1892, p. 144–147)

Kraft-Ebing says that in sexual bondage flagellation is "means to make coitus, or at least ejaculation, possible", but is not clear about what it is in flagellation that makes these possible. Is it the experience of a pain in itself that makes ejaculation possible? It is not obvious how this might be. As opposed to what Kraft-Ebing describes as masochism, it is not clear that in sexual bondage the pain experience plays an instrumental role. It could be that what is making the ejaculation possible is not the pain sensation itself, strictly speaking, but the arousal produced by the flagellation.

This lack of precision about the role of the pain experience is shared with the contemporary psychiatric definitions of masochism. According to these, an unpleasant pain might be desired and pursued as a means to obtain sexual pleasure. Masochists, it is claimed, use pain, and other typically unpleasant practices like being bound, beaten, asphyxiated, and humiliated, as means to sexual satisfaction.[29] However, these approaches to masochism do not clarify how pain plays an instrumental role in achieving sexual satisfaction. In fact, I think that in these cases the pain experience is not a means for sexual satisfaction, rather, they should be explained as examples of side effect pain. I will come back to these cases to argue that they do not seem to involve pains that are not unpleasant and, contrary to what might be thought, the pain experiences involved are not a means, they do not play an instrumental role for sexual pleasure. Whereas in masochism, as understood by Kraft-Ebing, the pain experience does seem to have an instrumental role for achieving the feeling of submission, it is not clear how *the feeling* of pain has an instrumental role in order to achieve sexual satisfaction in sexual bondage.

There is another example where people seek out pains, and these experiences do have an instrumental role. Goldstein (1983) offers this case. He argues that masochism is better understood as an example of irrational and abnormal self-punishment; masochism consists in using an unpleasant pain as a means for undeserved self-punishment. In his words:

> [T]here is *some* pain and unpleasantness which the masochist does want, namely, that which he deliberately inflicts… At the moment when Lise is slamming the door on her finger she is desiring *that pain*. Though there might be some pleasure and happiness which she does desire, namely, that which she

[29] Two good examples of these contemporary accounts of masochism are the Diagnostic and Statistical Manual of Mental Disorders (DSM-V), and the 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10).

> feels upon punishing herself… Philosophers sometimes think of the masochist as a calculating hedonist… Though there may be strains of reason in masochism, the concept of masochism entails irrationality and abnormality. (Goldstein, 1983, p. 223)

I take it that Goldstein thinks that it is in virtue of pain's unpleasantness, and the fact that unpleasantness is bad, that pain may be a means for punishment. If pain were not unpleasant and bad, it would not be a means for punishment. The fact that pain is unpleasant has an instrumental role for constituting a way of punishment, including self-punishment. Moreover, there might also be an element of pleasure involved, since punishment may relieve guilt and this might feel good. In either case, the unpleasantness of pain plays an instrumental role for punishment, regardless of such punishment additionally entailing some form of pleasure through relief. Goldstein argues that not all cases of self-punishment that involve pain are masochistic, though. He thinks that the difference between a normal case of self-punishment and masochism is that the latter is *irrational* self-punishment. Masochism is the infliction of an unpleasant pain as a means for self-punishment, where there is no good reason to be punished. If, for example, you punish yourself for something that you did not do by hurting yourself, this would count as masochism according to Goldstein.

All in all, an unpleasant pain might be pursued because we think of it as means, i.e., because we think that the unpleasant pain plays an instrumental role for getting something else that we want. Since we believe that an unpleasant pain is a means for something else that we want, that unpleasant experience is desired, which explains why people might seek out an unpleasant experience. In this way we can explain the *motivation* to engage in an unpleasant pain even if it is unpleasant: we might be motivated to pursue an unpleasant pain because we believe that such unpleasant pain will bring about something else that we desire, it is in virtue of desiring a pain as a means that we can explain why we sometimes seek out an unpleasant pain.

We can also be justified to seek out an unpleasant pain. That is to say, there might be some circumstances where it may be desirable to engage in an unpleasant pain as a means. For instance it could be that, all things considered, the final outcome of engaging in an unpleasant pain is good enough to compensate for its badness. I do not think that we need to enter into the details of what would constitute an end that is good enough to justify an unpleasant means. However, this can help us to understand how it might be worth seeking out unpleasant experiences as a means. If the final outcome is good enough, the end may

justify the means, even when the means is an unpleasant pain. I think that the crucial feature in means-end masochism is that people intentionally pursue an unpleasant pain because such experience is desired as a means for something else, whether or not there are good reasons to seek out that unpleasant pains as a means.

6.2 Side effect pains

There is a different type of case that might be easily mistaken for means-end masochism. People might willingly and knowingly put up with unpleasant pains, but without the experience really being a means. Rather, the pain is sometimes a side effect of what is really being pursued. We are aware that certain activities will result in a pain experience, yet we still engage in such activities; however, after more careful analysis, we can notice that pain is not playing an instrumental role. This may appear strange; why would someone engage in an activity that entails an unpleasant pain if unpleasantness is bad? And why would one do so if the unpleasantness of the experience is not a means to an end? The explanation is that, sometimes, unpleasant experiences are a side effect of other things that we might pursue as a means or as an end. That is to say, we sometimes engage in activities that involve unpleasant pains because they are an unavoidable side effect of something that we intentionally pursue. I think that Armstrong is an example of a theorist running different cases together:[30]

> It is true that we may *endure* pain, some even enduring it quite gladly, for the sake of something else. But this does not imply that we have a favourable attitude to the pain itself… [T]he case of the masochist, and also the type of neurotic who 'seeks punishment', may be raised here ... What happens, I think, is that he finds certain features of some situations so pleasurable that he *willingly puts up with the pain for the sake of that pleasure*. If this is correct, then *we all act like masochists when we deliberately swallow scalding hot tea in order to warm ourselves up*, or get into what is initially a hot bath. All these things involve a good deal of 'hedonic sophistication': we have to brace ourselves to some degree to do these things for the ultimate pleasure they bring. (Armstrong, 1962, p. 90–91, my emphasis)

[30] See Trigg's (1970, p. 157-162) discussion on the "enjoyment of pain" for a list of cases where, I think, means-end masochism is run together with side effect pains.

When Armstrong considers the cases of the "neurotic who 'seeks punishment'", this looks quite similar to what Goldstein had in mind: one might put up with an unpleasant pain because the unpleasant pain is taken to be a means. In contrast, when Armstrong talks about putting up with an unpleasant experience produced by drinking scalding hot tea, this does not imply that the unpleasant experience in itself is a means to an end. More precisely, an unpleasant pain may be an unavoidable consequence of the real means. The unpleasantness of this experience is playing no instrumental role in achieving the pursued end, in contrast to means-end masochism. When one puts up with the scalding water that burns one's tongue in order to warm up, it is not in virtue of one's tongue being burned, or of that burning experience being unpleasant, that one gets warmer; one gets warmer because the water that one drinks is hot, which has some bodily consequences that result in one feeling warmer. The fact that drinking scalding water might entail an unpleasant experience is a side effect of the means in order to get warm through drinking such hot water. Burning one's tongue and this being unpleasant is not a means; it is a side effect of the real means, where the end is to get warmer.

Many other cases might look as if some unpleasant pain, or its unpleasantness in particular, is a means to an end, but it is rather a side effect of the actual means. Going to the dentist would be a clear example of something like this. We do not go to the dentist *because* it hurts, we go *even if* it hurts. Most medical procedures that cause unpleasant pain take this form, I think. For example, you might need an injection in order get analgesics and, as a consequence, feel better. You really dislike injections because they hurt, but, at the same time, you feel really bad and finally get the injection that produces an unpleasant pain. In a sense, you want the unpleasant pain produced by the injection, that is, you are willing to put up with the unpleasant pain in order to feel better, which does not mean that feeling this unpleasant pain plays an instrumental role in making you feel better. In other words, if the injection were painless, it would still constitute the means to its end. If the unpleasantness of the experience produced by the injection were needed for feeling better, then such unpleasantness would be a means, but this is not the case. Similarly, we are willing to put up with the unpleasant pains of many medical procedures because such procedures are the means for something else. However, it is not in virtue of these procedures being painful that we get better. The same happens with an unpleasant tooth extraction with the dentist. It is not because the extraction hurts that we get the extraction. The unpleasant pain is only a side effect of a means.

There is another kind of case in which we might put up with an unpleasant pain as a side effect: the unpleasant experience might be a side effect of *an end*. We might be willing to

put up with an unpleasant experience because it is the result of something else that we want. However, in this case, the thing that we want might not necessarily be a means. For example, suppose that you want to celebrate that you finished your PhD thesis by drinking a whole bottle of champagne. Let us suppose that drinking a bottle of champagne is, as a way of celebration, an end in itself. However, you know that you always get a really bad hangover and your head hurts intensely when you drink alcohol. You are really joyful and finally decide to drink the whole bottle for celebration's sake. Why would you do this if you know it is going to be unpleasant to have a really bad headache afterwards? A good way to explain this is that when you take the decision, you are willing to put up with the consequences of your end. There are, indeed, many things that we want and for which we are willing to put up with the possible unwanted side effects, where unpleasant pains are one of those side effects.

We can explain the motivations behind these two different types of cases. We could shed light on why we engage in activities that involve unpleasant pains by appealing to our desires about our means or our ends, even if we know there are unavoidable unpleasant consequences of our means and of our ends. Moreover, it might be sometimes sensible to put up with the unpleasantness of a pain in order to achieve certain means or certain ends. It seems that it is often rational to go to the dentist, for instance. When we go to the dentist, the benefits in terms of health outweigh the badness of the unpleasantness that we might have to endure, and the same can be said for many medical procedures.

Going back to the psychiatric definitions of masochism and of sexual bondage, it is not obvious in these cases how the experience of pain plays an instrumental role in order to achieve sexual satisfaction. If these so called masochists would prefer to experience the same sexual satisfaction *without* having to experience pain, this would show that masochism as understood by the medical definitions is rather a case of side effect pain. If we do not consider that going to the dentist is masochism, we might want to stop considering people who pursue certain activities that involve pain as a means for sexual satisfaction as masochists. However, we might think that there is a fundamental difference between these scenarios. One might think that *only* masochism entails an element of irrationality, or of abnormality. Again, we could want to draw a distinction between side effect pain cases, like going to the dentist, and 'real masochism', which implies irrationality or abnormality. We might have the intuition that there is something abnormal about masochism, whereas it is very common and even desirable to put up with certain unpleasant experiences, like when we go to the dentist or undergo other medical procedures. We might have this intuition because the term 'masochism' implies some sort

of deviation, which probably comes from its psychiatric origins. Some cases of side effect pains might involve irrationality, of course, and we might want to identify these as 'masochistic'. I can think of two clear cases.

First, it would be irrational if we could achieve an end without having to put up with the unpleasantness of a means, yet we chose to experience that means, which involves an unpleasant experience. This type of scenario would be equivalent to going to the dentist for a tooth extraction and choosing to have no sedation, even if this would not bring about any benefit. Second, it would also be irrational if the goodness of the end were not good enough to compensate for the badness of the means, yet we choose to go through the unpleasant consequences of the means. For example, it would not be desirable to have a tooth extracted with no analgesia, in order to take care of a minor decay. We might think that the cases of sexual masochism, such as the ones considered by psychiatry, involve irrationality, and that is why only these are masochistic. In contrast, going to the dentist normally involves good reasons, which explains why it is not masochistic.

I think that we should not worry too much about which cases deserve to be called 'masochism', as long as we notice that there is a distinction to be drawn. The question of how we should call each of these cases is terminological and, even if it might be important, it is not the main focus of this chapter. In some occasions we put up with the unpleasantness of experiences such as a pain because it is the side effect of either a means or an end. We might be motivated to put up with such unpleasantness because we desire the means or the end, even if we think that the unpleasant experience is an unavoidable consequence of our means or of our ends. Further, there are occasions where we might have good reasons to put up with such unpleasantness, and in other cases we might not have good reasons for enduring it, it is not desirable. I think that it is intuitive to call 'masochistic' the situations where we seek out pain, whether we have good reasons to do so or not. In this sense, if someone is putting up with a pain as a side effect in order to get sexual satisfaction, and even if there were no good reasons for doing it, that person is not seeking out the pain experience. In this way, I do not consider these cases as masochistic, since pain is not sought out, strictly speaking.

In conclusion, none of these cases seem to suggest the existence of pain experiences that are not unpleasant. In all side effect cases, the fact that we are willing to put up with pain shows that these experiences are unpleasant. Side effect pains are not sought out at all, so they do not seem to be examples of purely pleasant pain, i.e., cases of non-unpleasant

pains. In contrast, in the next section I will consider cases of pain experience that are sought out as an end.

6.3 End-in-itself masochism

Even if many cases that involve unpleasant pains can be explained by showing that the unpleasant experience is not what is pursued in itself, there are other instances that call this into question. Numerous examples suggest that people sometimes desire, and intentionally pursue, unpleasant pain experiences as an end in itself. But how could it be that someone wants to experience something such as pain as an end, if pain is unpleasant, if pain is bad in itself? Could there be instances of pains that are pleasant? I think there are different ways in which pain can be an end in itself. These are two that appear in the philosophical literature about masochism: i) an unpleasant pain can be a part of a pursued whole, or ii) an unpleasant pain might be *also* pleasant. I will now develop each of these possibilities.

6.3.1 Contextual masochism

Some philosophers have noticed that, sometimes, something unpleasant might be a part of a whole that is sought. This is different from the previous cases. The unpleasant experience is not a means to something else or a side effect; instead, it is a constitutive part of a whole, where the whole is sought. The pain does not play an instrumental role in order to obtain something distinct, the pain is not a means. The pain is not a side effect either; it is not something that one is willing to put up with, and that has no instrumental role, in order to obtain something else. Instead, the pain is a part of a whole, and it is the whole that is desired as an end in itself. Pitcher (1970), for example, thinks the following about masochism:

> It is important to note that the masochist *does not just put up with the pain*, enduring it merely because the rest of the scene is so wonderful: on that account, we would all be masochists when we had a tooth removed or drank some ghastly medicine. No, the situation would lose its appeal for the masochist if the pain were to be removed from it. (Pitcher, 1970, p. 484, my emphasis)

I think that there are a few points to make about what Pitcher says. First, "putting up with the pain" may mean two different things: i) it could be that the pain is a side effect, where the pain does not have an instrumental role, or ii) it could be that the pain is sought as a means, where the pain plays an instrumental role. It is not obvious that Pitcher notices these two different cases. That said, he has arguably identified yet another kind of case where people seek out a pain experience. People want to feel pain because, otherwise, "the situation would lose its appeal". I think that Pitcher uses the term 'masochism' in a different way than the previous authors, and I think he offers a good example of the type of scenario that would count as contextual masochism:

> The notion of "liking (or disliking) something in itself" or of "finding something pleasant (or unpleasant) in itself" is a tricky one, and not at all so clear as it might superficially look to be. Suppose Agnew heartily dislikes olives, but is fond of paellas that contain them. And we are to imagine not that he barely tolerates the presence of the loathed olives in his paella; so that he loves the dish *despite* them: on the contrary, we are to suppose that he regards the olives as constituting a *necessary ingredient* in a really good paella — without them, it would be far less interesting. (Pitcher, 1970, p. 484, my emphasis)

There are a few things that I think should be said about Pitcher's example. First, paellas don't really have olives! But more importantly, this example provides us with an idea of what it may mean to seek out, desire, like, engage in, etc. a pain experience when it comes to contextual masochism. That is to say, in this scenario Agnew dislikes olives *in themselves*, but he likes them when they are *part of* paella. Why does Agnew only like olives when they are part of paella? This could be interpreted in different ways. It could be that olives taste different to Agnew when eaten as part of paella. It could be that now they have a pleasant taste. This could be possible. However, I think that Pitcher's intuition is that whether or not olives' flavour changes in some respect for Agnew, the important thing is that the taste of olives remains *unpleasant*.

The taste of olives is unpleasant for Agnew when he eats them on their own, in isolation from other flavours; he dislikes olives because their taste is unpleasant. However, when Agnew eats olives in paella, that olives' taste, together with its unpleasantness, is worth being pursued when it is accompanied by a myriad of other flavours, textures, aromas, etc. that occur when Agnew eats a paella. When the taste of olives is isolated, it is unpleasant and not worth pursuing for Agnew; nonetheless, it is an essential ingredient of good

paella, according to him. This does not mean that the olives' taste is no longer unpleasant for Agnew when he eats them in paella. Rather, that same unpleasant taste experience is an essential part of the whole paella-eating episode. By 'episode' I mean a series of events, such as different phenomenal experiences, that constitute a whole. The olives in paella maintain some of their unpleasantness, which explains why without them, eating paella would be far less interesting for Agnew.

Similarly, other cases where people engage in unpleasant experiences can be analysed as contextual masochism. For example, someone may find the whole episode of running a marathon worth pursuing. This is not so in the sense that the whole marathon episode is a distinctive sensory pleasure, but in the sense that there is something about the whole of running a marathon that is worth pursuing. This whole episode has, and must have, certain periods of intense unpleasant pain. The moments when one experiences an unpleasant pain are essential for the whole marathon episode to be worth pursuing: if running a marathon was not hard at some stages, it would not be worth engaging in it for some people. Let us illustrate this.

Suppose that you are running a marathon and you feel a strong pain in your leg at kilometre thirty-eight. This unpleasant pain, and particularly its felt unpleasantness, is not a means to finish the marathon. It is not in virtue of this unpleasantness that you can get the experience of running a full marathon. The unpleasantness of this pain is not tolerated as a side effect either. In this particular case, you do not want to run a whole marathon in spite of having *that pain*. Rather, for it to be worth you running a marathon, you must overcome certain difficulties such as to continue running despite having unpleasant experiences. The felt unpleasantness of pain at kilometre thirty-eight is constitutive, according to yourself, of the worthwhile whole marathon episode. You would not want to have such unpleasant pain experience in isolation, but in this context the unpleasant pain is embraced and is essential to the whole. This does not mean that such pain is not unpleasant or bad in itself, but that this pain is wanted as an end because it is constitutive of the marathon episode, which is sought out as an end.

Let us have a look into another culinary example. This one, in contrast with the olives case, is an example that involves an unpleasant pain. Contextual masochism may also occur with the experience produced by eating spicy food. In the same way that some dishes are partially constituted by spicy chilly peppers, the experience resulting from eating these dishes is partially constituted by the burning feeling in our mouths as we eat the dish. The burning feeling in our mouths that we sometimes feel when we eat spicy

food is caused by the capsaicin that some chilly peppers contain. This ingredient is partially constitutive of chilly peppers and, therefore, it is also partially constitutive of some spicy dishes.[31] Some spicy dishes are partially constituted, therefore, by ingredients that will cause one to get the feeling of a burn in one's mouth. The burning feeling is not pleasant in itself. The burning feeling caused by spicy chillies is actually unpleasant, but it is also a constitutive part of the whole episode of tasting certain spicy dishes. Again, in this example the burning feeling in one's mouth is not the means to obtain some form of pleasure, nor does one put up with the burning feeling just in order to taste the other flavours of the dish. Instead, the burning feeling is a constitutive aspect of the whole spicy food episode. The whole episode of tasting a spicy dish includes an unpleasant burning pain as one of its essential parts and, in those terms, such unpleasant experience is pursued as an end.

Once we take this into account we can explain one way in which we might purse unpleasant pains as an end. We might be motivated to engage in an experience that is partially unpleasant because, all things considered, we believe that the whole episode is worth pursuing and desire the episode as a whole. It is in virtue of our beliefs and desires about the whole that we can render intelligible one's behaviour. If the unpleasant pain felt while running a marathon were taken out of this context, we would not engage in such an experience. But given that it is a constitutive part of an end, we also pursue that unpleasant pain as an end. If the end as a whole is actually worth pursuing, that is, if the end is good enough, we will then also be justified in engaging in something that is partially composed of an unpleasant experience. If the whole is in fact good enough, this would not only explain why we might engage in these unpleasant experiences, but it would also constitute a good reason for pursuing the whole as an end.

There might also be circumstances where it might not be worth pursuing the whole. There might be situations of irrational contextual masochism. If we eat a spicy dish partially because of the burning feeling, or if we run a marathon partially because of the unpleasant parts within it, we still might have a good reason to engage and pursue these activities because the overall outcome is good enough to compensate the badness that it involves. These kinds of behaviour are quite frequent. We might think that there might be, in contrast, an irrational version of contextual masochism: the goodness of the overall outcome is not heavy enough to tip the balance to the positive side. In these cases, we might think that there are no good reasons to actually seek out activities that involve

[31] For more on the effects of capsaicin see Wood (1993).

unpleasant experiences. I think this might be possible. There might not be good enough reasons for some people to run a marathon. That said, whether it is rational or not, in cases of contextual masochism people seek out unpleasant pains that are essential parts of a whole.

So does the fact that we seek out pain in these contexts show that there are non-unpleasant pains? These pains are not sought as a means, so we might think that these pains are not unpleasant. However, the fact that pain is not sought as an end *simpliciter*, but as an essential part of a broader end, suggests that the pains involved in contextual masochism are unpleasant. There seems to be no reason to doubt that these pains, like the ones that we have when eating spicy food and while running marathons, are not unpleasant. There might be different reasons for seeking out these pains, but the reason does not seem to be that they are only pleasant. In the next section I will consider a case where pains are claimed to be pleasant.

6.3.2 Masochistic pleasures

Klein (2014) argues for another type of masochistic scenario. This is another possible type of scenario in which we might seek out a pain experience. According to Klein, some token experiences such as pain can be both pleasant and unpleasant. Klein calls these *masochist pleasures*. These pleasures might not be sexual and don't have to be pathological or abnormal. He proposes that these masochist pleasures might include experiences such as the ones we have when getting body modifications (piercing and tattoos), performing hardcore sports (like running a marathon), pleasantly dwelling on one's anger or jealousy, wiggling a loose tooth, enjoying the aesthetic ambivalence of a horror film, eating spicy food, being whipped (both in sexual and non-sexual contexts), etc. Klein thinks that masochistic pleasures, including pain, are pleasant *in addition to* being unpleasant and *because* they are unpleasant. In his words:

> To make this more concrete, suppose someone takes pleasure in being spanked. The current account says that there will be three distinct features that jointly characterize the experience. There is a first-order sensory state, the bodily pain, that arises from the spanking itself. That sensory state is *painful*. It hurts. That quality of painfulness is pleasant. The distinctive contribution of the penumbral account is to explain just when and why painfulness can be pleasant. (Klein, 2014, p. 52)

It is worth pointing out that Klein's terminology is slightly different from the one I have been using. What he calls "first order sensory state" is equivalent to the phenomenal experience of being a pain, what he calls "painfulness" is what I refer to as being unpleasant, and what he calls "pleasant" remains the same in my terminology. According to Klein, a pain can be a masochist pleasure when the pain is just within the boundary of being bearable. This is the *penumbral theory*. When an unpleasant experience is just within the limits of being bearable, the unpleasantness of such experiences may become additionally pleasant.

There are different reasons that explain why the unpleasantness of pain can also be pleasant in these situations, Klein thinks. For example, the unpleasantness of a barely bearable pain experience can become pleasant: i) if it is *novel*, ii) if it allows us to exercise *self-control*, iii) if it stands for *special intimacy* with someone, iv) if it *pushes our own boundaries* helping us to grow and change, etc. For example, when one is wiggling a loose tooth and it is a novel experience, then the unpleasantness of this experience can become additionally pleasant; "finding the edge of unbearability might be pleasant precisely *because* it is a surprising and novel discovery." (Klein, 2014, p. 50, my emphasis) If someone is being whipped and this causes an unpleasant pain, but it stands for a special intimate relation or pushes one's own boundaries, the unpleasantness of the pain caused by the whipping can become additionally pleasant. These experiences are unpleasant, yet they are pursued for themselves, insofar as the unpleasantness of these pains is also pleasant.

There are a few things that I think we must stress about Klein's view. First, masochistic pleasure is not contextual masochism. Klein insists that we should not misunderstand his proposal and think pleasantness is predicating the whole, which is partially constituted by an unpleasant pain. If this were the case, then masochistic pleasures would actually be a case of contextual masochism, i.e., we would judge the whole experience, that includes an unpleasant pain as a part, as pleasant and thus worth pursuing. Instead, Klein insists that self-control, novelty, intimacy, etc., together with the fact that an unpleasant pain is within the limits of being bearable, are the reasons in virtue of which only the unpleasantness of a pain might become additionally pleasant.

It is also important not to confuse Klein's proposal with means-ends masochism. That is, one does not pursue a masochistic pleasure because such experience will bring about something else that is pleasurable or pursued as a further end. Instead, some form of pleasure partially constitutes the masochistic pleasures, i.e., a masochistic pleasure is

composed by being unpleasant and by that unpleasantness being pleasant. One might be tempted to think that you wiggle your tooth and seek out pain because this gives you a pleasurable "sense of novelty" or a pleasurable "sense of control". This might be possible, but it is not the type of case that Klein considers a masochistic pleasure, which seem to be possible too. Finally, the unpleasant pain is not, clearly I think, a side effect of something else that is pursued, since the pain is sought out, rather than being something collateral.

There are a few last remarks about masochistic pleasures that I think are important. First, we could have the opposite situation, that is, the pleasantness of an experience being additionally unpleasant. I will not go into the details of these cases because it is not very clear which experiences could be an example of this; more importantly, these would still be unpleasant pains, and we are looking for instances of non-unpleasant pains. Second, Klein thinks that masochistic pleasures are very common, which shows that not everybody thinks that masochism implies some form of abnormality. Again, what I take to be the commonality in all cases of masochism is that people seek out pain experiences. Masochistic pleasures illuminate a different reason why we might seek out pains, i.e., we might seek out pains that are unpleasant if their unpleasantness is also pleasant.

It is because of this pleasantness that we can explain why we might be motivated and even have a good reason to engage in such masochistic pleasures. Additionally, masochistic pleasures may also include other aspects that explain why it may be worth pursuing them. For example, in order to have a masochistic pleasure we might also exercise self-control, stand in special intimacy with someone, or push our own boundaries thus helping us to grow personally. If seeking out pain experiences in masochistic pleasure is good enough, all things considered, we might be justified in pursuing masochistic pleasures, even if they entail unpleasantness. If the goodness of a masochistic pleasure does not outweigh the badness of the unpleasantness that it entails, all things considered, then we would not be justified in engaging in such masochistic pleasure. Again, the crucial feature of masochistic pleasures is that the unpleasantness of some experiences is additionally pleasant.

So, are there any instances of non-unpleasant pains in masochism? If we had found a case of masochism where one seeks out pains that are *only pleasant*, we could have concluded with no doubt that being a pain does not entail being unpleasant. However, I have not found any examples that point towards this direction. I do not think that there is clear evidence of people for who pain is only pleasant. If this were the case, they would act in extremely odd ways and they would be very unlikely to survive. Maybe some people

could experience a few pains that are only pleasant. Are there examples of this? I cannot think of any crystal clear cases of purely pleasant pains. Moreover, since pain is paradigmatically unpleasant, I think that all pains experiences that we have are, in fact, unpleasant.

6.4 Conclusion

In this chapter I have explained a wide diversity of possible masochistic scenarios that appear in the philosophical and psychiatric literature. The unifying feature among all of them is that they include an experience that is unpleasant, i.e., an experience that is bad in itself. We can explain and account for the reasons in virtue of which one might seek out such unpleasant experiences. Moreover, we can also explain the situations in which we might even be justified to engage in unpleasant pain experiences. Once we analyse the different circumstances, we can dissolve the apparent contradiction. We can make sense of the different reasons why we might pursue an unpleasant pain. The analysis of all these masochistic scenarios sheds light on our interactions with pain and unpleasantness. Moreover, it also illuminates the diverse reasons why we might seek out something that is bad in itself.

Even if there are no concrete instances of non-unpleasant pains, not even in masochism, it seems at least to be possible that there could be. The fact that we can conceive pain as not being hedonic, as we did during the last chapter, and that pain can be purely pleasant, suggests that we should have a theory of unpleasantness that allows this possibility. In the next chapter I will propose a theory that is able to capture this possibility, while also being able to account for the heterogeneity of unpleasant experiences, and having other advantages in contrast to the content and desire theories.

CHAPTER 7: THE HEDONIC DETERMINABLE-DETERMINATE THEORY

7.0 Introduction

In this chapter I propose a theory that has many virtues. This view maintains the strong intuition that unpleasantness is something felt, something phenomenal; it can account for the possibility of non-unpleasant pains; it can account for the *ploner case*, i.e., an example for which it is claimed that an unpleasant experience lacks only its pain phenomenal aspect; it can easily explain how two sensory experiences can vary only hedonically, i.e., one is pleasant and the other unpleasant; and finally, this theory is able to account for the heterogeneity of unpleasant experiences. This is the hedonic determinable-determinate theory (HDDT). In order to explain and defend this theory, I will proceed as follows.

I will contrast HDDT with the developed version of Crisp's proposal from Chapter Four. I will show how HDDT is a superior theory in many respects. The main difference between these theories is that in HDDT, *unpleasantness* is understood as a determinable property that has multiple determinates, i.e., multiple ways of being unpleasant that vary non-additively; however, being a pain is not one of these ways of being unpleasant. A sensory unpleasant experience such as an unpleasant pain is a mental state composed by two phenomenal properties: i) the phenomenal property of being a pain, and ii) the phenomenal property of being unpleasant *in a certain way* u1, u2, u3, etc. I will explain in detail what this means.

However, HDDT faces two important difficulties. First, it needs to account for the phenomenal variation among different ways of being unpleasant. In other words, if all unpleasant pains belong to the same kind, how can we explain that they feel so different from one another? I think that HDDT is able to capture this feature. For this, we need to explain which might be the *dimensions* along which different ways of being unpleasant may vary. The general idea is that in the same way that different ways of being coloured may vary along its dimensions (e.g., hue, saturation, and brightness), unpleasantness also has essential dimensions. The variation along these dimensions is what accounts for the variations among different ways of being unpleasant, different *unpleasantnesses*.

Finally, I will consider a second problem for HDDT. How can we account for the intrinsic badness of unpleasantness? To put it simply, it is not obvious why the phenomenal property of unpleasantness entails badness. HDDT does not offer a reductive account for the unpleasantness of pain, and does not explain why unpleasantness entails badness.

However, I will argue that even if HDDT does not offer an explanation of why an unpleasant experience is bad in itself, it stands in an overall better position than the theories discussed in previous chapters regarding the nature of unpleasantness. All in all, HDDT is the best available account for understanding the unpleasantness of pain.

7.1 The hedonic determinable-determinate theory (HDDT)

Let us quickly remember the last proposal from Chapter Four. This theory was able to account for the heterogeneity of unpleasantness using the determinable-determinate distinction. According to this view, being a pain is understood as a determinate of being unpleasant. This means that being a pain is a way of being unpleasant. Different phenomenal properties, e.g., being a pain and being an itch vary non-additively from each other. This means that they differ along the same dimensions, whichever these might be. This explained how this theory could account for the heterogeneity of unpleasant experiences, as I explained in Chapter Four. The fundamental problem with this view is that being a pain entails being unpleasant. As we saw in the previous chapters, there are no conclusive cases of non-unpleasant pains. However, the discussion about these cases shows that even if there are no actual cases, it is possible that there might be. It seems at least conceivable that pain experiences could exist without being unpleasant. In fact, some authors accept pain asymbolia as an actual example of this possibility. If it is possible to have non-unpleasant pains, a theory of unpleasantness should be able to capture this possibility.

I think that there is a straightforward way of dealing with this problem; this will be fundamental in offering a solution to this and many other problems concerning unpleasantness. If one uses the determinable-determinate distinction in order to account for the heterogeneity of hedonic experience, one should dissociate the property of being a pain from the property of being unpleasant. That is to say, we should apply the determinable-determinate distinction to the unpleasantness of pain, but being a pain should not be understood a way of being unpleasant. Instead, an unpleasant pain involves a certain way of being unpleasant. This is what HDDT proposes to account for unpleasant pains experiences.

The hedonic determinable-determinate theory for the unpleasantness of pain

An unpleasant pain experience has two phenomenal properties: i) the phenomenal property of being a pain, and ii) a phenomenal determinate property (u1, u2, u3, etc.) of the unpleasantness determinable.[32]

In this way, we can easily account for the *possibility* of asymbolics having pain experiences that are not hedonic at all. That is, they would have a mental state with the phenomenal property of being a pain, but without the phenomenal property of being unpleasant. Even if we accept that the evidence is not as conclusive as we might have thought to declare without doubt that asymbolics have non-unpleasant pains, the discussion about asymbolics is evidence of the possibility of non-unpleasant pains. If we accept HDDT's proposal for the unpleasantness of pain, we can accommodate this possibility. Similarly, in this way we could account for the possibility of pains being only pleasant, this is a possible form of masochism where people might seek out pain as an end. Even if there is no obvious evidence of instances of masochism along these lines, we can now account for the possibility. We could have an experience that instantiates the phenomenal property of being a pain, and the phenomenal property of being pleasant in a certain way pl1, pl2, pl3, etc., that vary non-additively from one another.

Another advantage of dissociating the property of being a pain from the property of being unpleasant is that we can account for cases where, presumably, someone has an experience that is unpleasant, without being a pain or a phenomenal experience in any other way. Some think that we could have phenomenal experiences that are just unpleasant, and that there is an actual example of this. For instance, some researchers claim that their "results demonstrate, for the first time in humans, a loss of pain sensation with preserved pain affect." (Ploner, Freund, & Schnitzler, 1999, p. 211) This is based on a single case of a man who, after a stroke, lost to a good extent the capacity to experience sensory experiences, such as the ones produced by thermal stimuli. According to this

[32] I should thank Dr. Jennifer Corns for various aspects of this proposal that come from her unpublished work, "Hedonic Independence", and other drafts of related papers in progress, as well as due to personal communication. That is, she defends that: i) there is at least one distinct hedonic quality space (i.e. distinct from any sensory, or even cognitive, quality spaces); ii) that each such hedonic quality space may be empirically discovered and constructed through "just noticeable" qualitative judgements, as I will explain in more detail in Section 7.2.1; and iii) that the unpleasant quality of each unpleasant experience is constituted by the values it takes for each dimension in the relevant hedonic quality space.

study, "[i]n the patient reported here, clinical examination and cutaneous laser stimulation revealed… *loss of sensory discriminative pain component and preserved motivational-affective dimension of pain*." (Ploner, Freund, & Schnitzler, 1999, p. 213) This case does not put into question the existence of non-unpleasant pains, these are not pain experiences after all, it is claimed. However, if this case is real, or even possible, HDDT allows us to incorporate it.

HDDT has another advantage: we can easily explain how different sensory experiences may vary *only hedonically*. We can explain how the same type of sensory experience can sometimes be pleasant and at other times unpleasant. Take, for instance, the gustatory experience of tasting chocolate; this experience is one of tasting chocolate in virtue of a phenomenal property chocolate-taste-feeling-ness, say.[33] We can have chocolate-taste-feeling experiences that are sometimes pleasant, and at other times unpleasant. How can we explain this? These two experiences, the one that is pleasant and the one that is unpleasant, seem to be ways of being chocolate-taste-feeling experiences. However, it does not seem that we can account for these experiences being different by appealing to the determinable-determinate distinction. It looks like the pleasant chocolate-taste-feeling experience and the unpleasant chocolate-taste-feeling experience vary *additively*. These experiences are chocolate-taste-feeling *plus* something else, i.e., they are either pleasant or unpleasant.

The solution is quite simple once we understand hedonic sensory experiences as being composed of distinct phenomenal properties. The pleasant experience of tasting chocolate is composed of two properties: i) the chocolate-taste-feeling-ness property, and ii) a determinate way of being pleasant that I will call *chocolate-pleasantness*. Then, the unpleasant experience of tasting chocolate is composed of two properties: i) the chocolate-taste-feeling-ness property, and ii) a determinate way of being unpleasant that I will call *chocolate-unpleasantness*. Two different sensory experiences that only vary hedonically are not different determinates of a common determinable; this is so because they vary additively. However, the pleasantness, or unpleasantness, of these particular pleasant and unpleasant sensory experiences, are determinate properties. For example, chocolate-unpleasantness is a determinate of unpleasantness, it is a way to be unpleasant. Unpleasantness is a determinable phenomenal property that has certain determinates such as chocolate-unpleasantness —although one could also call these ways of being

[33] I call the phenomenal property of having an experience of eating something that tastes like chocolate chocolate-taste-feeling-ness, as opposed to the property chocolate-taste-ness. Whereas the former is a property of mental states, the latter is a property of things that taste like chocolate, such as a cake.

unpleasant u1, u2, u3, etc. Something very similar applies to chocolate-pleasantness, which is a determinate property of the determinable pleasantness, but one could also call these ways of being pleasant pl1, pl2, pl3, etc. This, I think, should be enough to understand how two different sensory experiences can vary only hedonically, i.e., when one is pleasant and the other unpleasant.[34]

Finally, we should explain how HDDT deals with the heterogeneity of unpleasant experiences. The heterogeneity problem consisted in explaining how unpleasant experiences all feel unpleasant, yet they feel unpleasant in very different ways. We did not seem to be able to account for all the variations in the ways of being unpleasant by appealing to a unitary feeling of unpleasantness, as we saw in Chapter Three, not even when this feeling has different intensities, as we saw in Chapter Four. According to HDDT, unpleasantness is phenomenal, but there is no unitary feeling of unpleasantness because different unpleasantnesses vary non-additively. HDDT is in line with the intuition behind the heterogeneity problem.

More precisely, following HDDT for unpleasantness, all unpleasant experiences qualify as unpleasant in virtue of instantiating unpleasantness as a determinable property. All unpleasant experiences are unpleasant in a certain way, u1, u2, u3, etc. and each of these ways of being unpleasant, each of these unpleasantnesses, varies from the others non-additively. In other words, u1, u2, u3, etc., vary along the same essential dimensions.[35] Once we take into account that being unpleasant is a determinable property, we can explain how unpleasant pains vary in terms of their unpleasantness, even if there is not a unitary unpleasant-feeling that is shared among all unpleasant pains. HDDT maintains the intuition that unpleasantness is a phenomenal property, while successfully dealing with the heterogeneity of unpleasantness. In the same way that there might be different ways of being red, without all of them sharing a unitary redness plus something else, there are different way of being unpleasant, without all of these ways of being unpleasant sharing some unitary unpleasantness.

---

[34] I leave open whether pleasantness and unpleasantness are determinates of a common determinable. This will depend on whether these properties vary non-additively, i.e., if pleasantness and unpleasantness share the same essential dimensions of variation. In either case, I will consider in further detail what might be the dimensions of variation of unpleasantnesses in the next section of this chapter.

[35] One could argue that there are also different ways of being a pain, and that being a pain is also a determinable property. That is, that there are different ways of being a pain experience that are determinates of a common determinable. If this were correct, this would imply that there are different determinates of being a pain, pa1, pa2, pa3, etc. These would be phenomenal properties and would vary from one another non-additively. This could be a way to deal with heterogeneity concerning the pain aspect of certain experiences. However, I do not think that we need to get into the details of this in order to deal with the heterogeneity problem, since it focuses only on the unpleasantness of experiences.

Experiences can be hedonic in various ways; there might be different ways of being pleasant or of being unpleasant, and these differences are not only regarding their intensity. I am not the first to allow for the possibility that there might be various ways of being unpleasant. Labukt (2012), for example, argues in favour of a pluralistic version of hedonic tone.

> The view even looks fairly attractive. It retains the phenomenological plausibility of the hedonic tone approach… and avoids the difficulties of the alternatives, while at the same time being immune to the heterogeneity objection. Now we should not feel particularly confident that strong pluralism [i.e., the claim that there are different ways to be pleasant or unpleasant] is correct until we have some fairly well-founded ideas about what the different hedonic tones are. (Labukt, 2012, p. 199)

The good news is that I think we do have a fairly well-founded idea of what these hedonic tones are: they are phenomenal determinates, they are different ways of being unpleasant, or pleasant, and they vary from each other non-additively. While Labukt acknowledged that we should allow the possibility of there being different ways of being unpleasant in order to give an answer to the heterogeneity problem, HDDT provides the details of what these different ways of being hedonic are: they are phenomenal properties that should be understood based on the determinable-determinate distinction.

7.2 Two final problems for HDDT

Even if HDDT looks like a promising theory to understand the structure of unpleasant experiences, there are still some details that need to be filled in. I will focus on two important aspects of the theory that I think need to be developed before we can conclude that it is the best available candidate to account for the unpleasantness of pain. First, it is important to explain how *different* ways of being unpleasant belong to *the same* kind, i.e., how different determinates of unpleasantness may vary non-additively from one another. In the same way that we can account for different triangles belonging to the same kind, or different shades of red all being determinates of redness, we should be able to account for the different ways of being unpleasant. Second, we also need to say something about two very important aspects of unpleasantness: the fact unpleasantness is motivational and normative in itself. It is in virtue of an experience being unpleasant that we have reasons, and even good reasons, to perform various actions in relation to the unpleasant experience.

I will show that even if HDDT does not have an explanation for why this is so, it stands in a better position, all things considered, *vis-à-vis* its competitors from previous chapters.

7.2.1 The different ways of being unpleasant

HDDT is capable of solving the heterogeneity problem. However, we have not yet explained what could account for the variations among the different ways of being unpleasant. In order to give a more substantial account of HDDT, we need to clarify how we could know which are the *essential dimensions* of unpleasantness. This would explain how unpleasant experiences share a common phenomenal property, even if there is so much diversity regarding the different ways of being unpleasant. I will not provide the specific essential dimensions of variation for unpleasantness, but I will motivate the idea that unpleasantness should be understood as a determinable and that it must have essential dimensions that we can identify. I will proceed as follows.

First, I explain how other determinates of a common determinable can share the same dimensions. Second, I will show how the notion of a dimension can also be used to account for differences among phenomenal qualities. I will focus on how we can create a *quality space* based on the essential dimensions of colour experience. The quality space of the experience of colour is a map of the different dimensions that constitute the experience of being coloured. Finally, I will suggest that we could also create a quality space for different ways of being unpleasant. The quality space for unpleasantness would illuminate how different determinates of the unpleasantness determinable could vary from each other non-additively.

Let us start with the essential dimensions of being triangular. In what do different ways of being triangular consist in? In the case of the determinable property being triangular, or triangularity, one way of accounting for the different determinates of triangularity is to refer to the variations along the three essential angles of a shape. If there is a variation in one of these angles, we obtain a different way of being triangular. If there is no variation along these angles, there is no variation in the way of being triangular. Each way of being triangular is a determinate of triangularity, and each different determinate varies from the others non-additively. The difference in ways of being triangular does not consist in being triangular plus something else. Rather, all the different ways of being triangular are the results of variations along the same dimensions, i.e., variations in relation to the three angles that compose a shape.

One important feature of understanding triangularity as a determinable is that it allows us to see how we can provide different accounts of the dimensions of a determinable. That is to say, there might be other ways of explaining the variations among different ways of being triangular. For instance, we could account for different ways of being triangular by referring to the length of the three essential sides of a shape. In this case, we specify that being triangular means to be a shape composed by three sides, and the variations along the length of these sides accounts for various determinates of triangularity being different. The difference between being equilateral, isosceles, and scalene is based on variations along the lengths of the three essential sides of a shape. Being equilateral means to be a shape with all three sides of the same length, being isosceles means to have two sides of the same length, and being scalene consists in having all sides with different lengths. So, even if referring to the change in angles is a more accurate way of specifying different ways of being triangular, this shows that there are different ways of specifying the essential dimensions of a determinable, i.e., there might be different ways of specifying what it means for determinates of a common determinable to vary non-additively.

The fact that we can account for the essential dimensions of a determinable in different ways is important. It is relevant because this means that we do not have to commit to just one possible way of describing the dimensions of variation in order to show that a determinable must have certain dimensions. Further, even if we have an account for the dimensions of a determinable such as being triangular, these dimensions do not have to be the correct ones, or the only possible description of these dimensions, in order to show that being triangular is a determinable property. Whatever the dimensions of a determinable might be, it is variations along these dimensions that explains the variations among determinates of a common determinable. The fact that we can describe different ways of being belonging to a common group by referring to variations along *the same* dimensions, strongly suggests that these different ways of being are determinate properties of a common determinable. I think it is relatively clear which might be the dimensions of triangularity, but how can we find out which are the dimensions of phenomenal properties?

In order to answer this, I think we can appeal to the notion of quality space. The idea is the following. We have a good conception of the dimensions along which determinates of being coloured may vary from one another, i.e., hue, saturation, and brightness. When we considered the determinable coloured-feeling-ness, the dimensions of this determinable are surely phenomenal derivatives of the dimensions of being coloured, i.e., something like felt hue, felt saturation, and felt brightness. Rosenthal's (2015) notion of quality

spaces seems to vindicate this idea. He focuses on the quality space of the experience of colour. In order to build this quality space, subjects are confronted with different colour samples which vary in their hue, saturation, or brightness, and they are asked to distinguish the *just noticeable differences* between the samples. The structure of a quality space is built on these distinctions. The creation of the quality space is purely based on the judgments that people make regarding the just noticeable differences within their experiences of colour. Modifying the samples across the dimensions of colour generates the changes in colour perception in the subjects, and the quality space is built.

Different samples are instantiating different ways of being coloured and, the idea goes, different changes in perception are instantiating different phenomenal properties of ways of being coloured, differences regarding coloured-feeling-ness. The quality space is not meant to account for the differences in the colour of the samples, but to reflect the experienced differences. In other words, the quality space of colour is reflecting the *experienced* hue, saturation, and brightness. The creation of a colour quality space unveils the dimensions of coloured-feeling-ness, i.e., the dimension along which the coloured-feeling determinates may vary: experienced hue, experience brightness, and experienced saturation. The difference between a colour experience of Coca-Cola red and of Ferrari red is that these experiences are instantiating different phenomenal properties; the difference between these phenomenal properties is explained in terms of a variation along their experienced hue, saturation, and brightness.

According to Rosenthal (2015) we can apply the same methodology to different sensory modalities. However, it might be much more difficult to create quality spaces for other sensory modalities. Colour experience is relatively easy for creating a qualitative space, since the experience of being coloured seems to have very few dimensions that are rather easy to control for experimentation. Rosenthal thinks that we can create quality spaces for other sensory modalities, at least in principle. For example, the same methodology has been used to try to create a quality space for olfactory experiences (Young, Keller, & Rosenthal, 2014). It is much more difficult to create an olfactory quality space because there are many more possible combinations of molecules that can result in noticeable olfactory differences. The olfactory experience determinable, or olfactory-feeling-ness, would have many more felt dimensions and it would be much harder to make well-controlled samples that mirrored and instantiated these felt dimensions. Nevertheless, we could in principle create a quality space by providing subjects with a wide range of combinations of molecules to smell, and we could map the variations among different ways of being an olfactory experience. If we could create a quality space for olfaction, it

would not be a map of the different molecules that produce smell experiences; instead, it would be a model that would mirror the different dimensions along which the different olfactory determinates may vary.

Going back to unpleasantness, I think that there are in fact different ways of being unpleasant, different unpleasantnesses, and that, in principle, these variations could be explained by appealing to changes in their essential felt dimensions. The way in which the feeling of nausea is unpleasant seems to be quite different from the way in which the experience of a cramp in one's leg feels unpleasant. Similarly, the way in which feeling a numb limb is unpleasant is quite different from the way in which having a migraine is unpleasant. This is what we should read as lying behind the intuition of the heterogeneity problem. How can we explain that all these experiences are unpleasant, yet feel so different? We could think that unpleasantness is not a feeling, as desire theories suggest in Chapter Two, and explain the phenomenal differences without appealing to different ways of feeling unpleasant. However, if we think that unpleasantness must be phenomenal, and we also think that that there is no unitary way of being unpleasant, then we must accept that there are different ways of being unpleasant, various unpleasantnesses — even if we find ourselves incapable of delineating at a first glance the precise categories and types of ways of being unpleasant.

I think we could apply a similar strategy to build a quality space for unpleasantness, at least in principle. The quality space of unpleasantness would show the essential dimensions of this phenomenal determinable property, and would help us identify different unpleasantnesses. We could vary the different stimuli that normally cause the unpleasantness of unpleasant experiences and map the noticeable differences produced by these different stimuli. The quality space of unpleasantness will not be a map of the different stimuli that normally cause the unpleasantness of unpleasant experiences. Rather, it would be a map of the different experienced unpleasantnesses. Moreover, the variations in stimuli that produce the unpleasantness of experiences would translate into noticeable differences of unpleasantness, they would show us which are the felt dimensions of unpleasantness. Similar to the felt dimensions of smell experiences, the felt dimensions of unpleasantness might be much more complex than the three-dimensional model that we can create for colour experience. However, the space quality model of unpleasantness could be created, at least in principle.

As we saw in Chapter Four, *intensity* might be one of these dimensions. That is to say, a mild unpleasant headache and an excruciating unpleasant headache vary in the way they

are unpleasant. Intensity is, I believe, one of the dimensions along which different determinates of unpleasantness may vary. However, intensity would not be the only dimension of unpleasantness, since there are differences in the ways of being unpleasant that cannot be reduced to variations in intensity. Even if the model that would map the variations of different ways of being unpleasant may be quite complex, I do not see why this could not be done in principle. So, what could these precise dimensions of unpleasantness be?

I do not know which would be exactly the kind of stimuli that could result in just noticeable differences regarding the ways in which an experience can be unpleasant. However, for illustration sake, let us accept that being *dull* and being *sharp* are different ways in which a headache might be unpleasant. If this were correct, the differences between being dull and sharp are not variations in the sensory component of an unpleasant pain, but differences regarding their unpleasantness. If we accept this, we could try to find the stimuli responsible for the noticeable differences between being dull and being sharp, and then this would allow us to identify the shared dimensions along which dullness and sharpness may vary.

Similarly, if intensity is a dimension of unpleasantness, we could identify the stimuli responsible for increases in the intensity of unpleasantness in order to create the quality space of unpleasantness. Even if this is an oversimplification of what the dimensions of a quality space of unpleasantness might look like, I do think that this is enough to make a more relevant point: that given that there are different ways of being unpleasant, and that these different ways could be explained by variations along shared dimensions, this confirms that unpleasantness is a determinable property. In the same way that we can create a quality space for the experience of colour, we could do so for unpleasantness. Just by trying to create a quality space of unpleasantness, we could come up with a more refined vocabulary about the different ways of being unpleasant u1, u2, u3, etc. Experimenting with different ways of being unpleasant, I think, would also shed light on the dimensions along which these different unpleasantnesses may vary from one another non-additively.

Even if we do not know *exactly* which are the dimensions along which unpleasantnesses may vary, we can accept that unpleasantness' dimensions exist. At the end, I think it is quite easy to accept that many different unpleasant experiences do feel different *qua* their unpleasantness. In an analogy, just as we can appreciate and accept the variation of different visual or olfactory experiences, even if we know nothing about their precise

dimensions of variations, we can accept that there are different ways of being unpleasant, without yet knowing which are the precise dimensions along which unpleasantnesses might vary. As Funkhouser puts it:

> [T]he concept of a determination dimension is quite general and it applies to various kinds — kinds with essences that are *phenomenological*, functional, *qualitative*, etc. *We can certainly disagree over the determination dimension* of a particular kind, but so long as this disagreement is reasonable, *the very existence of such disagreement helps confirm that we share an intuitive understanding of the concept of determination dimension and the task of discovering them.* Determination dimensions are simply those *essential dimensions of a kind along which instances of that kind can vary*. Scientific kinds are of particular interest to the metaphysicians and I assume that the determination dimensions for such kinds typically are to be discovered by *a posteriori* investigation. In particular, the science of a given kind should provide us with the determination dimension for that kind. (Funkhouser, 2014, p. 30, my emphasis)

If Funkhouser is right, it is probably an empirical affair to discover which are the precise felt dimensions of unpleasantness as a kind, in the same way that it is rather an empirical affair to discover the dimensions along which being an experience of colour, or of odour, may vary. However, it is a philosophical affair, it seems to me, to point out that the determinable-determinate distinction helps us to understand the structure of unpleasantness in order to address the heterogeneity of unpleasantness. In conclusion, HDDT is the best candidate to account for the heterogeneity of unpleasantness if we think that unpleasantness is phenomenal. Furthermore, this approach opens the possibility of an empirical research that explains the variability among the large diversity of different ways of being unpleasant. However, HDDT has an important limitation shared with most other theories of the unpleasantness of pain, as we will see in the next section.

7.2.2 The intrinsic badness of unpleasantness

I will now move on to the final important problem for HDDT, and for any other theory that takes an unpleasant pain to be motivational and normative in virtue of its unpleasantness. This is particularly problematic for the theories that take unpleasantness to be something felt, that is, that an unpleasant pain is motivational and normative in virtue of feeling unpleasant. The problem is to explain why unpleasantness entails non-

instrumental badness, and how if unpleasantness is phenomenal, this phenomenal property entails such badness. It is by appealing to this intrinsic badness that we can explain why some felt experiences motivate and justify actions merely in virtue of feeling the way they do, while other phenomenal experiences do not motivate or justify action in virtue of feeling the way they do. For example, a headache seems to motivate and justify certain actions because of how it feels, in virtue of feeling unpleasant, i.e., because having an unpleasant headache is bad in itself because it feels bad. In contrast, other phenomenal experiences are not bad in themselves; the experience of seeing something red does not seem to motivate or justify action merely because of how this experience is phenomenal, this experience is not motivational or normative in virtue being red-feeling.

The intrinsic badness is an important yet usually unexplained feature of unpleasantness. It is important because being motivational and normative are two crucial features of unpleasant experiences such as unpleasant pains. The fact that unpleasant pains are bad in themselves gives them a special position regarding action in contrast to other phenomenal experiences. Furthermore, the fact that unpleasant pains are normative in virtue of being bad makes these experiences very important in relation to moral issues. For instance, it seems that at least one of the reasons why it is morally relevant to consider pain infliction as a form of torture is because unpleasant pains are bad in themselves. We care about which creatures experience unpleasant pains, and thus have at least one form of moral relevance, also because unpleasant experiences seem to be bad in themselves. In contrast, there are other phenomenal experiences that do not instantiate badness, and these kind of phenomenal mental states are not morally relevant in themselves.

I will not offer an account of why unpleasantness entails intrinsic badness. If we think unpleasantness is phenomenal, I will not explain why only certain phenomenal properties entail badness, and why unpleasant experiences are motivational and normative in virtue of such intrinsic badness. I take it as a brute fact that unpleasantness entails non-instrumental badness. The experiences that instantiate unpleasantness have motivational and normative force because having one of these experiences is bad in itself. Hedonic experiences that are pleasant or unpleasant are both motivational and normative. That is, it because of certain experiences being hedonic that we can render certain related actions intelligible, that these hedonic experience are motivational, and that we can also justify related actions, i.e., hedonic experiences may constitute good or bad reason for actions, hedonic experiences are also normative.

Most authors, including myself, take as a given that unpleasantness entails non-instrumental badness. It seems clearly true that unpleasant experiences are bad in themselves. However, none of the considered theories about unpleasantness provides a developed explanation of why having an unpleasant experience is bad in itself. I take the badness of unpleasantness to be primitive, i.e., there seems to be no more fundamental explanation of why this is so. All theories that take the non-instrumental badness of unpleasantness as a primitive do not explain what accounts for unpleasantness to entail badness. The problem with most content theories is not that they accept this badness as a given.[36] The problem for content theories is that they fail to account for unpleasantness' badness by appealing to the mental content that is supposed to constitute such unpleasantness. The main virtue of content theories was to illuminate unpleasantness by appealing to mental content, but mental content cannot shed light on the non-instrumental badness of unpleasantness. This is precisely what is shown by the messenger-shooting problem.

HDDT is similar to most content theories in the sense that it accepts that unpleasantness is phenomenal and bad in itself. Since experiences that are unpleasant *feel bad*, we can explain why unpleasantness is motivational and normative. Why did you take a painkiller for your headache? Because the headache was unpleasant, it felt bad; the unpleasant pain is motivational in virtue of being unpleasant. Do you have a good reason to take a painkiller? Yes, the fact that unpleasantness is bad in itself provides us with a good reason to take a painkiller and thus get rid of the unpleasant pain; the unpleasant pain is normative in virtue of being unpleasant. Both HDDT and content theories agree that unpleasant pains are motivational and normative because they are mental states that instantiate unpleasantness, and a mental state that is unpleasant is bad in itself. HDDT and most content theories are in an equal situation regarding the fact that they accept that felt unpleasantness entails badness. So how does HDDT do any better?

First, HDDT is preferable to content theories regarding the structure of unpleasantness, i.e., regarding the heterogeneity of unpleasantness. Content theories take all unpleasant pains to be unpleasant *in the same way,* as if unpleasantness was a unitary feeling, as if it was a clear ingredient that all and only unpleasant experiences have, including unpleasant pains. This is not the case, as the heterogeneity problem shows. HDDT is better because it

---

[36] Martínez (2015) does not think that unpleasantness is non-instrumentally bad. He thinks that unpleasantness is only instrumentally bad, which I think is counterintuitive, as I explained in Chapter One. When I mention content theories that take unpleasantness as non-instrumentally bad, I do not include Martínez' theory.

gives an account for the heterogeneity of unpleasantness. However, one may think, could content theories use HDDT as a blueprint in order to account for the heterogeneity of unpleasantness? One might think that content theories could accept HDDT in order to explain the diversity within unpleasant experiences and then explain such diverse phenomenology by appealing to mental content. Different mental content would account for different ways of being unpleasant, as long as these differences were explained non-additively.

However, I do not think that this strategy is available for content theories very clearly. Let us start with evaluativism, also referred to as BAP. According to this view, the determinable unpleasantness would consist in representing a bodily disturbance as being bad for oneself. What would the determinates of this determinable be? The determinates should consist in different ways of a bodily disturbance being represented as bad for oneself. However, if it is not very clear in this approach what it means for bodily disturbances to be bad, it is even less obvious what it means for bodily disturbances to be bad for oneself in many different ways; what would one's multiple badness consist in? I do not think that the HDDT strategy is clearly available for BAP.

Let us consider the two versions of imperativism. According MAP, the unpleasantness of pain consisted in the command "See to it that bodily damage *d does not exist*". If the determinable unpleasantness is constituted by this command, what would the determinates consist in? I do not think it is clear what it would mean to *not exist* in different ways. After all, according to MAP, the notion of unpleasantness seems to assume a unitary feeling of unpleasantness. Finally, according to KAP, the unpleasantness of pain consisted in a command about another sensation. The unpleasantness determinable would consist in the command "Don't have that sensation!". Which could be the determinates of this command? Are there different ways of *not having* a sensation? I think that content theories could, at least in principle, benefit from HDDT in order to explain the heterogeneity of unpleasantness. However, it is not very clear to me how they could *actually* implement HDDT in terms of mental content.

Second, HDDT does better than content theories regarding the nature of unpleasantness. The advantage to invoking evaluative or imperative content was to explain the phenomenology of unpleasantness, and all that comes with this, in terms of mental content. Among the things to be explained was the motivationality and normativity of felt unpleasantness. However, these theories fail to account for motivationality and normativity by appealing to mental content. So, the idea that unpleasantness consists in

such content is unmotivated. More importantly, one of the main purposes of content theories was to explain the intrinsic badness of unpleasantness by appealing to mental content. However, one cannot explain the intrinsic badness of an unpleasant pain in terms of something that is not bad, as the messenger-shooting problem shows. Content theories fail to accomplish one of their mains aims. As a result of this, it is better not to explain such intrinsic badness at all. HDDT does not offer a reductive account of unpleasantness, and, in not doing so, it does not have to commit to inconsistencies such as saying that a headache is unpleasant and bad in itself, but that we do not have non-instrumental reasons to take a painkiller to stop feeling the headache.

Let us now compare HDDT with the desire theories from Chapter Two. In this case, both HDDT and desire theories are able to account for the heterogeneity of unpleasant experiences. Desire theories can explain the great diversity of unpleasant experiences by pointing out that such unpleasantness is not actually something phenomenal, some unitary felt quality of experience, but merely the fact that certain phenomenal experiences are desired not to occur. Being unpleasant, according to desire theories, is not something felt. The fact that unpleasantness is not something felt allows them to avoid the heterogeneity problem: in this way they do not have to commit to the idea that unpleasantness is a unitary feeling. These theories still have to say something about unpleasantness entailing badness, since they also take unpleasantness to be bad in itself. In contrast to all the discussed theories, desire theories do not have to explain that it is *a feeling* that is bad in itself, because they do not take unpleasantness to be phenomenal. In fact, and to be fair, desire theories can say something about the intrinsic badness of unpleasantness by appealing to the intrinsic badness of desire frustration. If desire theories and HDDT can assess the heterogeneity of unpleasant experiences, and if desire theories can say something more about the intrinsic badness of unpleasantness, why should we prefer HDDT then? I think we should prefer it because it maintains the very strong intuition that unpleasantness is phenomenal, that it is something qualitative, something felt.

The strength of HDDT becomes more visible when we consider desire theories' Euthyphro dilemma, when we consider whether i) a pain is unpleasant in virtue of the fact that we desire not to have that pain, as desire theories postulate, or if ii) we desire not to have a pain because it is unpleasant, as HDDT suggests. More precisely, desire theories explain that what constitutes pain's unpleasantness is that we desire a hedonically neutral pain sensation not to occur. An unpleasant pain is unpleasant *because* we have a sid-desire

not to have a pain sensation.[37] Desire theories could *explain* why it is instrumentally good to have these sid-desires by referring to a teleological evolutionary explanation. However, as I explained in Chapter Two, there seems to be no motivating reason or non-instrumental reason to have sid-desires for hedonically neutral pain sensations not to occur, because, according to desire theories, there's nothing *intrinsically bad* about a hedonically neutral pain sensation.

In contrast, HDDT can give a rather straightforward justification for our desires in relation to pain and unpleasantness. According to HDDT, we are justified in desiring not to have an unpleasant pain because unpleasantness is bad in itself, because unpleasant pains *feel bad*. Unpleasantness in this view is bad in itself and phenomenal. Both HDDT and desire theories give an answer to the heterogeneity problem, but HDDT is preferable because it maintains the strong intuition that unpleasantness is something felt and thus can rationalize in a simple way the desires about pain experiences. We desire pain not to occur because pains are always in fact unpleasant, and because unpleasantness is a phenomenal property that entails badness.

Desire theorists could try to borrow HDDT and argue that the different ways of feeling unpleasant can be explained in terms of desires. They could try to argue that different ways of feeling unpleasant are constituted by different sensory experiences *plus* a sid-desire for that sensory experience not to occur. However, these different ways of feeling unpleasant would not be determinates of a common determinable because they would vary additively from one another. If desire theorists tried to argue that unpleasantness is phenomenal and that such phenomenology is somehow explained in terms of sid-desires, it seems that they would have to argue that there is a unitary feeling of unpleasantness, and that this unitary unpleasantness is constituted by the same kind of sid-desire. As I have argued, there seems to be no unitary feeling of unpleasantness among all and only unpleasant experiences. Furthermore, even if desire theorists accepted that there are different phenomenal ways of being unpleasant, it is not clear how different ways of being unpleasant could be explained in terms of variations of sid-desires, since a sid-desire is precisely *one way of desiring*. For all this, it seems like the best way of making sense of desire theories is to argue that unpleasantness is not phenomenal. However, HDDT allows us to deal with the heterogeneity of unpleasant experiences while maintaining the intuition that unpleasantness is phenomenal.

---

[37] Just as a reminder, in Chapter Two I explained what a sid-desire is, i.e., a simultaneous, intrinsic, *de re* desire about a sensory experience.

Finally, HDDT is in a much better position than any of the other theories that take unpleasantness as a phenomenal property without giving a reductive account of it. In comparison to the distinctive feeling theory, from Chapter Three, and the hedonic dimension theory, from Chapter Four, HDDT can offer an account for the heterogeneity of unpleasantness. The developed version of Crisp's proposal was also able to deal with the heterogeneity of unpleasant experiences. However, HDDT is in a better position regarding the relation between pain and unpleasantness. It is also capable to account for: i) the possibility of pains not being hedonic, ii) experiences being unpleasant without being sensory, as the ploner case suggests, and iii) different sensory experiences varying only hedonically. I conclude, therefore, that HDDT is the best available account to understand the nature of pain and unpleasantness.

7.3 Conclusion

In this chapter I have argued for HDDT as being the best available theory to account for the unpleasantness of pain. The proposal is, broadly, to understand that an unpleasant pain is constituted by two phenomenal properties: i) being a sensory pain, and ii) being unpleasant in a determinate way u1, u2, u3, etc. I think that the best way to understand unpleasantness is as a determinable phenomenal property that entails badness. This theory is able to account for the heterogeneity of unpleasant experiences, the fact that unpleasant experiences feel in a certain way, that unpleasantness motivates and justifies certain actions, and that it is possible to experience non-unpleasant pains, even if all the pain experiences that we have seem to be, in fact, unpleasant.

CONCLUSION

Let me end this journey by briefly restating each of the chapters.

In Chapter One, I argued that even if content theories provide an interesting reductive account of the phenomenology of the unpleasantness of pain, these theories fail to illuminate various aspects of unpleasantness. First, the representational *indicative* content theories fail to illuminate why the unpleasantness of pain is motivational. The *imperative* content views try to explain this feature in terms of imperative content, which is meant to be inherently motivational. However, content theories face a more fundamental problem. If the unpleasantness of pain is bad in itself, which explains why we are justified in doing something to stop feeling an unpleasant pain such as taking a painkiller, content theories cannot capture this feature by appealing to the content that unpleasantness is supposed to consist in. Moreover, if content theories accept that felt unpleasantness entails badness, and given that we have the very strong intuition that it does, then they are committed to saying that we do not have non-instrumental reasons to take a painkiller in virtue of felt badness, even if the felt unpleasantness of pain is bad in itself and should provide a non-instrumental reason for action.

In Chapter Two, I explained the desire theories for the unpleasantness of pain. According to this view, an unpleasant pain experience is unpleasant if and only if it is composed by two components: i) a pain sensation that is inherently hedonically neutral, and ii) a simultaneous, intrinsic, *de re* desire of that pain sensation that it not to be occurring, i.e., a sid-desire. This theory has some important positive features. It can illuminate, for instance, why unpleasantness is bad by referring to the badness of the frustration of the sid-desire that constitutes unpleasantness. This approach can also shed light on the heterogeneity of unpleasant experiences. It can explain that unpleasant experiences qualify as such even if there is no unitary feeling of unpleasantness that is shared among all and only unpleasant experiences. Instead, what unifies unpleasant experiences is that they are partially composed by a sid-desire. However, in doing this, desire theories abandon the intuition that unpleasantness is something felt. Moreover, they cannot account for a motivating reason or a non-instrumental justification for sid-desires about hedonically neutral pain sensations, since these pain sensations are not bad in themselves.

In Chapter Three, I accounted for the distinctive feeling theory. According to this approach unpleasantness is phenomenal, and it is a unitary feeling that is shared among all and only unpleasant experiences and by dint of which these experiences qualify as

unpleasant. This theory has the advantage over desires theories, I propose, that it can claim that unpleasantness is phenomenal and that it entails badness. Given that pains are unpleasant, this view provides a straightforward explanation for our desires in relation to pain. We desire pain experiences not to occur because they feel unpleasant. However, this theory faces a fundamental problem: the heterogeneity problem. After careful introspection, there seems to be no unitary feeling that is shared among all and only unpleasant experiences. I argued that the distinctive feeling theory is incapable of showing that there is such a unitary feeling of unpleasantness and, therefore, this theory should be rejected.

In Chapter Four, I analysed two more theories that take unpleasantness to be a phenomenal property. The difference between these and the distinctive feeling theory is that they appeal to the notion of dimension. First, the hedonic dimension theory compares the unpleasantness of an unpleasant pain to the volume of a sound. This allows us to account for the variations of different unpleasant experiences in terms of intensity. However, this is insufficient to account for the heterogeneity of unpleasant experiences, since the intensity of unpleasantness does not seem to be the only type of variation regarding unpleasantness. Second, we can make use of the determinable-determinate distinction and argue that being a pain is a determinate of the unpleasantness determinable. This means that being a pain is a way of being unpleasant, and different ways of being unpleasant vary from one another non-additively, i.e., that they are different because they vary along their shared essential dimensions. Unpleasantness in this view is a phenomenal property, but there is no unitary feeling of unpleasantness. According to this view, being a pain entails being unpleasant, just as being scarlet entails being red. However, there are actual cases that suggest that being a pain does not entail being unpleasant.

In Chapter Five, I turned to analysing the relation between being a pain and being unpleasant, by considering various cases that suggest that there are experiences of pain that are not hedonic, i.e., that are neither pleasant nor unpleasant. I divided these cases into two categories: i) pain insensitivity and ii) pain indifference. Pain asymbolia seems like the clearest example that involves pains that are not hedonic. In fact, some authors think that pain asymbolia confirms this. However, I show that there is no consensus about the interpretation of this condition regarding the existence of pains that are not hedonic. Moreover, I think that pain asymbolia could be interpreted as an instance of pain insensitivity. Given that pain is the paradigmatic example of an unpleasant pain, and since the evidence about the existence of pains that are not hedonic is not definitive, I concluded

that it seems like all pains that we actually have are hedonic. More precisely, all unpleasant experiences that we actually have seem to be unpleasant.

In Chapter Six, I considered a different possibility for pains that are not hedonic: masochism. The fact that people sometimes seek out pain suggests that these experiences are pleasant rather than unpleasant. If this were the case, this would show the existence of non-unpleasant pain experiences. I delineate different scenarios where people might intentionally pursue an activity that involves pain. First, in means-end masochism people might seek out an unpleasant pain experience because this experience is desired as a means to an end. Second, in side effect pain we might put up with pain experiences because they are the side effect of either our means or our ends. Third, in end-in-itself masochism people might seek out pain experiences because they are either an essential part of a whole that is desired as an end, or because pain is pleasant in addition to being unpleasant. Even if there is no conclusive evidence of pain experiences that are only pleasant, i.e., of non-unpleasant pains, we can shed light on the different reasons why we might seek out activities that involve unpleasant pains.

In Chapter Seven, I proposed a theory that is able to account for the possibility of non-unpleasant pains, while also illuminates the heterogeneity of unpleasant experiences. This is the hedonic determinable-determinate theory (HDDT). According to HDDT an unpleasant pain experience is composed of two phenomenal properties: i) the phenomenal property of being a pain, and ii) a phenomenal determinate property (u1, u2, u3, etc.) of the unpleasantness determinable. This theory, among other advantages, can account for the heterogeneity of the unpleasantness of unpleasant experiences without claiming that unpleasantness is a unitary feeling that is shared among all and only unpleasant experiences. Instead, unpleasantness is understood as a determinable with multiple determinates, different ways of being unpleasant, and the differences among each of these determinates can be explained in terms of variations along their shared essential dimensions, whichever these might be.

Given the foregoing chapters, I hope the reader finds that HDDT is the best available theory to account for the unpleasantness of pain. According to this approach, unpleasantness is a phenomenal property that entails badness. Unpleasant pain experiences feel bad, and this explains why we are motivated and justified in avoiding unpleasant pains. The fact that unpleasant pains feel bad gives us a reason, and a good reason, to take a painkiller. This account allows for the possibility of non-unpleasant pains, even if all pain experiences that we have seem to be unpleasant. This account also allows us to

explain how different experiences from various sensory modalities might be pleasant or unpleasant. Finally, this theory allows us to account for the heterogeneity of unpleasantness, without claiming that unpleasantness is a unitary feeling that all and only unpleasant experiences share.

LE CARACTERE DESAGREABLE DE LA DOULEUR

INTRODUCTION

La douleur est désagréable. Étant donné que la douleur est l'exemple paradigmatique d'une expérience désagréable, mon but est de clarifier ce qui caractérise la douleur et le caractère désagréable. J'essaie notamment de comprendre ce que peut signifier qu'une douleur soit désagréable et ainsi d'élucider la structure des expériences désagréables. Ce faisant, j'aborde plusieurs aspects problématiques de la relation entre la douleur et le caractère désagréable des expériences. Je fournis également une compréhension générale de ce que signifie pour une expérience non nécessairement douloureuse d'être désagréable. Voici quelques questions essentielles que j'aborde concernant la douleur et le caractère désagréable :

- En quoi consiste la douleur ?
- Que veut-on dire lorsque l'on dit que la douleur est désagréable ?
- Comment expliquer la diversité des expériences désagréables ?
- Comment expliquer que le caractère désagréable de la douleur puisse motiver une action ?
- Comment expliquer que le caractère désagréable de la douleur justifie une action ?
- Pourquoi recherchons-nous parfois des expériences douloureuses ?
- Une douleur est-elle toujours désagréable ?

Ma thèse répond à chacune de ces questions en détail et s'organise dans ce but en sept chapitres correspondant à trois problématiques principales : i) qu'est-ce qui constitue le caractère désagréable de la douleur ? (Chapitres 1 & 2), ii) comment rendre compte de la grande diversité phénoménale du caractère désagréable des expériences ? (Chapitres 3 & 4), et iii) dans quels cas la douleur n'est-elle pas désagréable ? (Chapitres 5 & 6). Dans le dernier chapitre (Chapitre 7), j'offre une réponse générale aux trois problématiques principales en proposant ma propre théorie sur le caractère désagréable de la douleur. D'après cette théorie, une expérience désagréable est une expérience ressentie, le caractère désagréable est défini comme une propriété phénoménale des états mentaux, et cette propriété doit être comprise en utilisant la distinction déterminable-déterminant. Mon travail fournit ainsi une compréhension détaillée de la nature de la douleur et du caractère désagréable.

CHAPITRE 1 : THÉORIES DU CONTENU

Dans ce premier chapitre, j'explique et analyse les théories selon lesquelles la notion de contenu mental peut expliquer le caractère désagréable de la douleur. La première section porte sur les théories dites de *contenu représentatif*. Deux approches distinctes sont détaillées : la théorie de la douleur de la représentation du dommage (DD) et la théorie de la douleur d'après Bain (DB), aussi connu sous le nom d'évaluativisme. La deuxième section de ce chapitre est consacrée aux théories dites de *contenu impératif*, aussi connu comme impérativisme. Deux approches distinctes sont détaillées : la théorie de la douleur d'après Matrínez (DM) et la théorie de la douleur d'après Klein (DK).[38]

Dans la troisième section du chapitre, je me concentrerai sur le problème principal de tout type de théorie du contenu. Toutes ces théories, peu importe leur version, doivent affronter un problème que je désigne comme « *meutre du messager* ». Les théories du contenu ne parviennent pas à clarifier une caractéristique cruciale de toute douleur désagréable en faisant appel au contenu qui est censé constituer ces expériences : une expérience douloureuse désagréable est mauvaise en soi et nous donne une bonne raison d'agir pour la faire stopper. C'est à cause de ce sentiment déplaisant que nous sommes justifiés de prendre des analgésiques, par exemple. Enfin, dans la quatrième section, j'étudie deux solutions possibles à l'objection de « *meutre du messager* ».

1.1 Théories de la représentation

Lorsque l'on pense à la douleur, on pense souvent aux dommages corporels. Par exemple, une définition standard de la douleur telle que celle fournie par l'Association Internationale de la Douleur (IASP) établit que la douleur est liée au dommage corporel. D'après une approche représentationnelle, une expérience douloureuse désagréable représente des lésions corporelles :

[38] Pour un développement de cette théorie concernant l'intensité de la douleur voir Klein et Martínez (à venir) , et pour la théorie impérative en termes de signaux de la douleur voir Martínez et Klein (2016).

Théorie de la douleur de la représentation du dommage (DD)

> Une douleur désagréable est une expérience mentale constituée par la représentation d'une partie de notre propre corps comme étant endommagée.

Une douleur désagréable est la représentation de notre propre corps comme étant endommagé.[39] Avoir une douleur et ressentir de la douleur comme telle consiste en la représentation d'une partie de notre propre corps comme étant endommagée, c'est-à-dire, la douleur est expliquée par la réception des informations de notre propre corps. Selon DD, si nous voulons expliquer pourquoi une douleur désagréable se ressent comme telle, par exemple, nous devrions faire appel au contenu de l'expérience. Ce contenu est ce que l'expérience de la douleur est censée être, et ce contenu devrait rendre compte de la phénoménologie d'une douleur désagréable. Cependant, il y a un problème fondamental pour DD. Cette théorie ne peut pas expliquer pourquoi les douleurs désagréables sont motivantes en soi, c'est-à-dire, pourquoi le contenu d'une expérience douloureuse désagréable peut nous motiver à faire quelque chose. En effet, il n'y a pas d'implication logique entre d'une part le fait de représenter quelque chose (de précis ou non) et d'autre part la motivation à faire quelque chose.

Il y a une solution assez simple pour DD. On peut ajouter un composant évaluatif au contenu qui constitue une douleur désagréable. C'est en vertu de ce contenu évaluatif qu'on pourrait expliquer ce qui rend une expérience désagréable et motivante. Cette approche est également appelée évaluativisme (Bain, 2012).[40]

La théorie de la douleur d'après Bain (DB)

> Le fait que le sujet ait une douleur désagréable consiste i) dans le fait que ce suject ait une expérience (la douleur) qui représente une perturbation d'un certain type et ii) que cette même perturbation soit représentée comme mauvaise pour soi-même dans le sens corporel.

Ce virage évaluatif est destiné à expliquer pourquoi les douleurs désagréables sont désagréables et motivent ainsi à agir. Une douleur désagréable ne représente pas seulement une information neutre qui pourrait être précise ou pas. Une douleur désagréable consiste en la représentation d'informations évaluatives. Autrement dit, une

---

[39] Voir aussi Tye (1995, 2006).

[40] Bain n'est pas le premier à défendre une forme d'évaluativisme. Voir Helm (2001, 2002), par exemple. Cependant, je considère que Bain a la version la plus claire et paradigmatique du évaluativisme.

douleur désagréable transmet des informations sur des faits évaluatifs, elle représente des objets comme étant mauvais. Selon DB, une douleur désagréable est la représentation d'une perturbation corporelle qui est représentée comme mauvaise pour soi-même. Le fait que la perturbation corporelle soit représentée comme mauvaise est censé expliquer pourquoi l'expérience fait mal, est désagréable, et pourquoi elle est motivante en vertu de faire mal.

Néanmoins, il y a un problème pour expliquer cette motivation en termes du contenu évaluatif. Ce que signifie pour un objet du monde d'être mauvais et, plus précisément, ce que signifie pour une partie de notre corps d'être mauvaise, n'est pas clair. Prendre en compte *l'argument de la question ouverte* de Moore (1903) permet de clarifier ce problème. Selon Moore, on ne peut pas donner une définition des concepts normatifs en se référant aux phénomènes purement naturels qui ne sont pas eux-mêmes normatifs. Les partisans de DB sont vulnerables à cette critique s'ils offrent une approche purement naturaliste du caractère désagréable de la douleur. Si nous acceptons l'argument de la question ouverte de Moore, alors DD et DB ont un problème assez similaire. Ces théories ne permettent pas d'expliquer comment des propriétés normatives, par exemple « être mauvais » pour, peuvent être motivantes en elles-mêmes, alors même qu'elles ne sont que la représentation de propriétés naturelles.

1.2 Théories impératives

Alors que les théories représentationalistes du contenu essayaient d'expliquer le caractère désagréable de la douleur simplement en termes indicatifs, en expliquant ce que représentent les expériences douloureuses désagréables, les théories imperativistes expliquent le caractère désagréable de la douleur à partir de ce que les expériences nous ordonnent de faire. Une bonne façon de souligner la différence entre ces deux types de contenu mental est de faire appel à la notion de *direction d'ajustement* (Anscombe, 1957; Searle, 1979). Selon la théorie impérativiste, nos expériences corporelles telles que la faim, la soif, les démangeaisons et les douleurs sont motivantes parce qu'elles sont constituées d'un contenu impératif (Hall, 2008). De la même façon que l'on pourrait penser que les désirs sont intrinsèquement motivationnels, étant donné qu'ils ont une direction d'ajustement du monde à l'esprit, si les douleurs désagréables sont conçues comme ayant une orientation similaire, alors les douleurs sont aussi intrinsèquement motivantes. Je propose ainsi une reformulation de la proposition de Martínez (2011, 2015):

La théorie de la douleur d'après Matrínez (DM)

Une douleur désagréable est constituée i) d'un contenu indicatif : «Il y a une perturbation corporelle *d* dans votre propre corps», ce qui explique que l'expérience soit une douleur et ii) un aspect hédonique du contenu indicatif « Veillez que la perturbation corporelle *d* n'existe pas! », ce qui explique le caractère désagréable.

Deux aspects de cette proposition sont particulièrement importants. Tout d'abord, en distinguant clairement les deux aspects d'une douleur désagréable, on fournit une explication simple de ce qui constitue le caractère motivationnel d'une douleur désagréable. Le fait qu'une douleur désagréable soit partiellement constituée par une commande explique pourquoi avoir une douleur désagréable est motivante. Deuxièmement, DM maintient l'un des principaux objectifs de Martínez, à savoir, le caractère désagréable de la douleur ne prescrit pas une action particulière. Le caractère désagréable de la douleur nous motive à agir, mais il y a beaucoup d'actions différentes qui pourraient être adéquates pour faire face à une expérience de la douleur. Cependant, DM doit également faire face à certaines difficultés. DM est censé expliquer ce qui constitue la dimension motivationnelle d'une douleur désagréable, mais un tel contenu impératif pourrait ne pas être suffisant en soi pour motiver l'action (Bain, 2011).

Klein (2007, 2012, 2015b) propose une approche différente de la douleur et du caractère désagréable. Deux aspects de sa théorie peuvent être utiles pour expliquer le caractère motivant d'une commande. Le commandement i) doit provenir d'une source d'autorité, dans ce cas-ci, notre propre corps, et ii) nous devons nous soucier de la commande, car le commandement provient d'une source d'autorité. Selon Klein, les commandements dans le cas d'une douleur sont émis par nos propres corps, et vise notre intérêt. Par ailleurs, selon lui, nous traitons notre corps comme une autorité pratique et nous nous soucions des ordres émis par celui-ci, ce qui permet d'expliquer qu'une commande nous motive à agir.

La théorie de la douleur d'après Klein (DK)

Une douleur désagréable est constituée par deux impératifs : i) un ordre « Gardez *B* de *E* (avec priorité *P*)! », ce qui explique que l'expérience soit une sensation de douleur, et ii) un autre ordre « N'ayez pas cette sensation de douleur! », ce qui constitue le caractère désagreable de l'expérience.

Ce contenu est destiné à rendre compte de tous les aspects de la phénoménologie d'une douleur désagréable. D'après DK, nous pouvons expliquer comment différentes expériences douloureuses désagréables sont ressenties en se référant à des variations dans le contenu du premier impératif. Il y a trois éléments principaux qui constituent une expérience de douleur : « *B* représente une partie en particulier du corps, *E* est une phrase gérondive passive nominalisée et *P* une fonction de classement»[41] (Klein, 2015b, p. 57). Lorsque vous ressentez de la douleur parce que vous vous êtes tordu la cheville, cette expérience consiste en un ordre « Gardez votre cheville de porter du poids ! », où cet ordre a une priorité par rapport à d'autres états mentaux motivationnels.

Ces trois éléments ne permettent de rendre compte que de l'aspect sensoriel d'une douleur. Le caractère désagréable de l'expérience s'explique en vertu d'un commandement concernant un autre ordre. En particulier, le caractère désagréable de la douleur consiste en un ordre qui demande de faire stopper une sensation de douleur spécifique. C'est une différence importante entre DM et DK. Alors que DM comprend l'aspect sensoriel de l'expérience composé par un contenu indicatif, DK explique le même aspect de l'expérience avec un contenu impératif. Néanmoins, un problème majeur se pose pour toutes ces théories du contenu.

1.3 Le problème de « meurtre du messager »

Jacobson (2013) a formulé ce qu'il a nommé l'objection de « meurtre du messager ». Quand un messager apporte de mauvaises nouvelles au roi, il n'est pas rationnel pour le roi de tuer le messager comme un moyen de faire face à la mauvaise nouvelle. Recevoir le message n'est pas en soi mauvais, ce qui est mauvais est ce dont le message informe le roi, en supposant que les mauvaises nouvelles sont correctes. Si le messager informe que les gens dans le royaume sont en train de mourir, le message n'est pas mauvais : il est mauvais que les gens meurent. Alors qu'il serait rationnel d'agir pour empêcher les gens de mourir, il serait illogique de tirer sur un messager qui apporte de mauvaises nouvelles dans le but d'agir sur le contenu de ces nouvelles. Le roi n'a aucune bonne raison de tuer le messager, puisque le fait que quelqu'un donne de mauvaises nouvelles n'est pas en soi mauvais.

DB est analogue à la situation du messager. Le simple fait de représenter quelque chose dans notre corps comme mauvais n'est pas mauvais en soi, car il n'est pas mauvais en soi

[41] Traduit de l'anglais.

de recevoir un message sur quelque chose de mauvais. Ce qui est susceptible d'être mauvais est ce qui se passe dans notre corps, dans un cas, et que les gens meurent dans le royaume, dans l'autre. De la même manière qu'il n'est pas rationnel (ni nécessaire, ni souhaitable) de tirer sur le messager, dans le cas de DB, il ne serait pas rationnel de cesser de représenter le mal corporel, de faire cesser cette représentation.

Or la théorie DB devrait pouvoir expliquer pourquoi il est en fait rationnel d'agir dans le but de faire cesser une sensation désagréable de douleur. Lorsque nous prenons un analgésique, c'est précisément ce que nous faisons : nous faisons cesser cette expérience. Si DB ne peut pas expliquer en quoi prendre un analgésique est rationnel, nous devons conclure que DB est incorrecte. Cet argument et ce problème du « meurtre du messager » se pose de la même façon pour les impérativistes. Recevoir un ordre n'étant ni bon ni mauvais en soi, les imperativistes ne peuvent pas non plus expliquer le caractère intrinsèquement mauvais d'une douleur désagréable. Ils sont également obligés de dire que la prise d'analgésiques n'est pas rationnelle.

1.4 Solutions au problème du meutre du messager

Pour répondre à ce problème, les imperativistes ont proposé qu'une douleur désagréable est mauvaise au sens où elle interfère avec d'autres activités et c'est à cause de cette interférence qu'il est rationnel d'agir par rapport à cette expérience, par exemple, de prendre un analgésique (Klein, 2015b; Klein & Martínez, à venir; Martínez, 2015). Selon les impérativistes, la prise d'analgésiques peut s'expliquer sans avoir à accepter que la douleur soit désagréable en soi. Au lieu de cela, c'est une raison instrumentale qui explique pourquoi nous prenons des analgésiques. Nous prenons un analgésique parce qu'une douleur désagréable interfère avec notre liste mentale de choses à faire.

Néanmoins, la torture semble être un contre-exemple. Supposons que l'on soit torturé par le moyen d'une douleur désagréable ; c'est une très mauvaise situation. Si on pouvait prendre une pilule qui nous permettrait d'éviter ces douleurs, on aurait de bonnes raisons de la prendre. Il semble incorrect de penser que la seule raison pour laquelle il est mauvais d'être torturé soit que la douleur désagréable que l'on experimente interfère avec d'autres choses que l'on aimerait faire. Ressentir une douleur désagréable est instrumentalement mauvais, sans que cela implique qu'une douleur désagréable ne soit pas aussi mauvaise en soi.[42] Je pense que le problème rencontré par les imperativistes pour rendre compte du

[42] Bain (2017) présente une intuition similaire.

caractère intrinsèquement mauvais des douleurs désagréables est structurel. L'impérativiste doit montrer qu'une douleur désagréable n'est pas mauvaise en soi, c'est-à-dire, que le caractère intrinsèquement mauvais consiste seulement en une raison instrumentale.

Les représentationalistes ont proposé deux stratégies principales afin de fournir une réponse au problème du meurtre du messager. Ces deux stratégies font appel à d'autres états mentaux. Il serait rationnel de prendre un analgésique en vertu de i) une émotion négative dirigée vers la douleur désagréable (Boswell, 2016) ou ii) parce que on a une aversion générale à la douleur désagréable (Cutter & Tye, 2014). La solution de Boswell repose sur des modèles empiriques de douleur selon lesquels une expérience normale de la douleur est constituée de trois éléments : i) une sensation de douleur brute (pas bonne ou mauvaise), ii) une composante affective, et iii) une composante émotionnelle négative.[43] D'après ce modèle, c'est en vertu de la composante émotionnelle de la douleur qu'il est rationnel de prendre un analgésique. Cette solution rencontre aussi des problèmes. Montrer qu'il y a des raisons instrumentales pour prendre un analgésique n'implique pas qu'il n'y a pas non plus de raisons non instrumentales.

Cutter et Tye expliquent qu'on a une raison de prendre des analgésiques parce que l'on a une aversion pour les douleurs désagréables. Cette aversion peut-être formulées en les termes d'une *con-attitude* générale. Autrement dit, le caractère mauvais non instrumental de la douleur désagréable peut être expliquée en faisant appel au désir de ne pas avoir une douleur désagréable. Ce n'est pas une forme d'instrumentalisme parce que nous n'expliquons pas le caractère mauvais d'une douleur désagréable à partir du caractère mauvais d'un désir. L'approche de Tye et Cutter est la meilleure solution proposée par les théories du contenu pour répondre au problème du meurtre du messager.

Il y a, néanmoins, au moins deux problèmes avec cette solution. Le premier est que si cette explication peut rendre compte de notre motivation pour prendre un analgésique, elle n'explique pas de façon claire la justification de cette action. L'aversion en elle-même n'implique pas une justification pour agir. Le deuxième problème est que cette théorie n'est plus capable d'expliquer en quoi le caractère mauvais d'une douleur désagréable réside dans sa phénoménologie. Autrement, il n'est plus possible d'expliquer le caractère mauvais de la douleur désagréable en utilisant la notion d'un contenu mental.

---

[43] Voir par exemple Fields (1999), Gracely (1992) et Price (2000).

CHAPITRE 2 : THÉORIES DU DÉSIR

Ce deuxième chapitre traite d'une autre manière de rendre compte du caractère degréable de la douleur. Cette approche prend appui sur les désirs afin d'expliquer ce qui constitue le caractère désagréable de la douleur et d'autres expériences déplaisantes (Armstrong, 1962; Brady, 2017; Pitcher, 1970). Je clarifierai d'abord ce que ces théories du désir entendent par désir. Ensuite, j'expliquerai comment ces théories utilisent la notion de désir pour rendre compte du caractère desagréable de la douleur. Je me concentrerai sur la proposition de Heathwood (2006, 2007, 2011). Enfin, une fois cette théorie du désir clarifiée, je soulignerai deux problèmes rencontrés par cette approche. Le premier problème a été identifié par Bramble (2013) ; selon lui les théories du désir ne peuvent pas expliquer ces cas où on a des expériences désagréables inconscientes. Le second problème, que je considère comme une objection fondamentale aux théories du désir, est souvent présenté sous la forme d'un dilemme d'Euthyphron. Je montrerai que bien que les théories du désir aient une réponse à ce dilemme, leur réponse reste problématique.

2.1 Le caractère désagréable de la douleur compris comme un désir

Une bonne façon de comprendre ce que sont les désirs est de les comparer aux croyances[44]. La différence entre ces deux types d'états mentaux est souvent expliquée en termes de direction d'ajustement, distinction que l'on doit à Anscombe (1957, §32).[45] Les croyances et les désirs ont une nature très différente : alors que les croyances ont une direction d'ajustement de la parole vers le monde, les désirs ont une direction d'ajustement du monde vers la parole. Les théories du désir ont une approche réductionniste du caractère désagréable. Les expériences désagréables sont constituées par deux états mentaux plus fondamentaux : i) une expérience sensorielle hédoniquement neutre, c'est-à-dire, une sensation qui n'est ni agréable ni désagréable, et ii) le désir de ne pas avoir cette expérience sensorielle (Heathwood, 2007). D'après les théories du désir, une expérience douloureuse désagréable est constituée par et ce réduit à ces deux éléments. Ils sont nécessaires et suffisants pour avoir une douleur désagréable.

[44] Pour plus des détails sur la notion du désir voir Schroeder (2017, 2004).

[45] Voir Searle (1979) pour la terminologie de *direction d'ajustement* [direction-of-fit].

Les théories du désir

> Une expérience de douleur désagréable est désagréable si et seulement si elle est constituée par deux composants : i) une sensation de douleur intrinsèquement neutre, et ii) un sid-désire, i.e., désir simultané, intrinsèque, *de re,* de ne pas avoir cette sensation de douleur.

L'un des principaux avantages des théories du désir est qu'elles peuvent expliquer l'hétérogénéité des expériences désagréables. Feldman (2004) fournit un bon exemple de cette intuition : « les plaisirs sensoriels sont tous des « sentiments », mais ils ne sont pas « ressentis de la même manière » » (Feldman, 2004, p. 79)[46]. Cette intuition se retrouve pour les expériences désagréables. Si vous pensez à plusieurs expériences désagréables que vous pourriez avoir, comme ressentir de la douleur, ressentir des étourdissements, ressentir des démangeaisons, de la faim, de la soif, etc., il n'y a rien de phénoménal, pas de sentiment unitaire conscient, en vertu duquel toutes et seulement ces expériences peuvent être regroupées comme appartenant au même type d'expérience. Par example, Korsgaard (1996) écrit : « Qu'est-ce que la nausée, la migraine, les crampes menstruelles, les coups d'épingle et les picotements ont en commun qui nous fait dire qu'elles sont douloureuses? » (Korsgaard, 1996, p. 148).

Le problème de l'hétérogénéité

> Nous avons une intuition très forte que : i) toutes les expériences désagréables et seulement ces expériences sont ressenties comme désagréables ; c'est en vertu d'être ressenties comme désagréable qu'elles qualifient en tant que telles. Cependant, après une introspection minutieuse, il y a aussi l'intuition forte que : ii) il n'y a rien de qualitatif, rien de phénoménal, pas de sentiment unitaire, que toutes les expériences désagréables partagent et en vertu duquel elles sont désagréables.

Je pense que la façon la plus simple et la plus élégante pour les théories du désir de confronter le problème de l'hétérogénéité est de dire que le caractère désagréable n'est pas qualitatif, que ce n'est pas quelque chose de phénoménal : le caractère désagréable n'est pas un sentiment. Plutôt, les théories du désir peuvent expliquer que ce qui unifie toutes les expériences désagréables, ou au moins toutes les expériences sensorielles désagréables, est qu'elles sont partiellement constituées par un sid-désir. Le problème avec cette solution est qu'on doit sacrifier l'intuition que le caractère désagréable est en effet quelque chose

[46] Toutes les traductions sont personnelles.

ressenti, que le caractère désagréable est un sentiment. L'avantage, par contre, est qu'on peut expliquer d'une manière simple ce qui unifie toutes les douleurs désagréables et toutes les expériences sensorielles désagréables.

Si on pense qu'un état mental est motivant si on peut expliquer un comportement en référant au tel état mental, alors nous pourrions expliquer comment la douleur désagréable est motivante. Pour cela, on compte sur le fait que les douleurs désagréables soient constituées par un désir. De plus, étant donné qu'il est mauvais en soi d'avoir des désirs frustrés et que la douleur désagréable est constituée par des désirs frustrés, alors il est mauvais en soi d'avoir des douleurs désagréables. Nous sommes donc justifiés à agir pour ne pas avoir de douleur désagréable et ainsi éviter d'avoir des désirs frustrés.

2.2 Problèmes pour les théories du désir

Bramble (2013) soulève un premier problème pour ces théories du désir. L'argument est que i) nous pouvons avoir des expériences désagréables inconscientes, mais ii) nous ne pouvons pas avoir des désirs sur des expériences inconscientes et, par conséquent, iii) ce qui explique le caractère désagréable de ces expériences inconscientes ne peut pas être un désir. Les théories du désir sont censées rendre compte des expériences désagréables, mais elles ne peuvent pas expliquer pourquoi des expériences inconscientes peuvent être désagréables. Haybron (2008) propose un exemple d'expérience sensorielle auditive inconsciente désagréable :

> Peut-être avez-vous vécu avec un réfrigérateur qui gémissait souvent à cause d'un mauvais roulement. Si c'est le cas, vous pourriez avoir constaté qu'avec le temps, vous avez complètement cessé de le remarquer. Mais parfois, lorsque le compresseur s'est arrêté, vous avez remarqué le silence soudain et glorieux ... Bref, vous aviez eu une expérience désagréable sans le savoir.[47] (Haybron, 2008, p. 222)

Même si nous acceptons que nous puissions avoir des expériences inconscientes désagréables, l'argument de Bramble n'est pas tout à fait convaincant. Il n'y a pas de raison de penser qu'on ne puisse pas avoir des désirs à propos de telles expériences désagréables, même inconscientes. En fait, ces désirs peuvent être soutenus soit au niveau conscient ou inconscient, c'est-à-dire que nous pourrions avoir conscience d'avoir ces

[47] Traduction de l'anglais.

désirs ou pas. Heathwood (à venir) a récemment offert une réponse similaire. Selon lui, le type d'exemple que Bramble a en tête implique qu'une sensation soit inconsciente en un sens restreint, et dans ce sens restreint ou faible, nous pouvons avoir des désirs à propos d'expériences dont nous sommes pas conscients.

Le deuxième problème des théories du désir prend la forme d'une version du dilemme d'Euthyphron : la douleur est-elle désagréable parce que nous désirons ne pas l'avoir, ou est-ce que nous désirons ne pas avoir de la douleur parce que celle-ci est désagréable ? En gros, alors que les théories du désir prennent la première option du dilemme, la deuxième option semble être beaucoup plus intuitive mais pas disponible pour les théories du désir. Les théories du désir offrent cette solution initiale au dilemme :

> Celui qui réduit le plaisir sensoriel au désir répond « Oui et Oui». Oui, nous désirons des sensations agréables parce qu'elles sont agréables (autrement dit, nous les désirons d'avance parce que nous savons que nous les désirerons quand nous les aurons). Et oui, les sensations agréables sont qualifiées comme agréables parce qu'elles sont intrinsèquement désirées.[48] (Heathwood, 2007, p. 39)

Le dilemme d'Euthyphro reste problématique. Heathwood veut répondre « oui » aux deux options du dilemme, mais il n'a pas vraiment pu dire « oui » aux deux. La réponse de Heathwood est insatisfaisante parce que, s'il peut effectivement expliquer les désirs de douleur dirigés aux douleurs désagréables, il ne peut pas expliquer les désirs simultanés dirigés aux douleurs qui sont hédoniquement neutres. Il n'a pas expliqué les raisons pour lesquelles on n'aurait pas envie de quelque chose qui n'a rien de mauvais en soi-même.

Il y a une autre stratégie disponible pour essayer de résoudre le dilemme d'Euthyphron. Cette stratégie consiste à dire que ce n'est pas la sensation de douleur qui possède la propriété d'être désagréable, mais plutôt le composé de la sensation de la douleur qui est hédoniquement neutre et du désir dirigé vers cette sensation de douleur (Brady, 2017). Cette solution offre une explication de ce qui constitue une douleur désagréable, sans impliquer que quelque chose qui n'est pas désagréable en soi devienne désagréable parce qu'on désire de ne pas l'avoir.

Si nous acceptons cette explication, il n'y a plus de dilemme. Néanmoins, il y a encore un problème non résolu. La proposition des théories du désir n'a pas encore pu rendre compte

---

[48] Traduction de l'anglais.

du désir même qui constitue une douleur désagréable. La solution proposée par Brady rencontre un problème similaire à celui rencontré par la solution de Heatwood : dans les deux cas, on ne peut expliquer pourquoi il est possible d'avoir des désirs envers une sensation de douleur hédoniquement neutre.

Les théories du désir peuvent proposer une explication évolutionniste à ce problème. Selon cette explication, le désir de ne pas avoir des sensations douloureuses hédonistiquement neutres fait partie d'un processus de sélection naturelle. Il n'y a rien d'intrinsèquement mauvais ou horrible dans les expériences de la douleur, mais nous avons évolué pour ne pas les désirer intrinsèquement ; « L'évolution a très bien fait son œuvre, et presque toutes les créatures vivantes du règne animal trouvent désagréables les sensations qui accompagnent presque tous les mécanismes de nociception. » (Hall, 1989, p. 648)

> **L'explication évolutionniste**
>
> Il est utile que des créatures comme nous désirons (avec un sid-désir) de ne pas avoir de sensations de douleur hédoniquement neutres, parce que de cette façon nous évitons les dommages corporels et nous sommes donc plus susceptibles de survivre et de rester en bonne santé.

Cependant, même si l'explication évolutionniste fournit une justification instrumentale des raisons pour lesquelles nous pourrions avoir des désirs de ne pas avoir des sensations de douleur hédoniquement neutres, elle demeure insuffisante : i) nous n'avons pas encore expliqué la raison qui nous motive à avoir ces désirs, et ii) nous n'avons pas fourni des raisons non instrumentales qui expliquent que l'on ait ces désirs. Ceci est problématique notamment si l'on considère comme fortement intuitif le fait que ce sont ces types de raison que la douleur offre par rapport au désir. Il semble particulièrement intuitif que les désirs concernant la douleur soient motivés et justifiés d'une façon qui n'est pas instrumentale, la douleur étant une raison motivante, une bonne raison de vouloir ne pas ressentir de la douleur. Or l'explication évolutionniste offre seulement une justification instrumentale de tels désirs et va à l'encontre de cette intuition.

Indépendamment des difficultés mentionnées ci-dessus, il y a une raison plus simple de rejeter les théories du désir. L'autre option du dilemme d'Euthyphro est toujours disponible. Pourquoi désirons-nous éviter des douleurs désagréables ? Parce que les douleurs désagréables sont ressenties comme désagréables et ce sentiment désagréable est mauvais en soi.

CHAPITRE 3 : LA THÉORIE DU SENTIMENT DISTINCTIF

Dans ce troisième chapitre, j'expose la théorie du sentiment distinctif. J'expliquerai le premier problème rencontré par cette théorie. Il s'agit d'expliquer comment le sentiment désagréable distinctif peut être une raison motivante et normative d'agir. Je défends l'idée que pour rendre compte du fait que le désagréable est motivant et normatif, on doit comprendre cette propriété comme mauvaise en soi. Le problème principal de cette théorie est donc un problème d'hétérogénéité. La théorie du sentiment distinctif établit que toutes les expériences désagréables sont désagréables en vertu d'une propriété phénoménale partagée qu'ont toutes les expériences désagréables. Cependant, selon le problème de l'hétérogénéité, il n'y aucun aspect qualitatif, il n'existe aucune propriété phénoménale unitaire parmi toutes les expériences désagréables, et en raison de laquelle ces expériences sont qualifiées comme désagréables.

3.1 La théorie du sentiment distinctif

Selon Bramble (2013), les expériences sensorielles agréables et désagréables sont respectivement agréables et désagréables en vertu d'un sentiment distinctif. Un sentiment distinctif est une qualité distinctive, c'est-à-dire un aspect phénoménal en vertu duquel une expérience sensorielle est qualifiée comme agréable ou désagréable. Une bonne façon de comprendre ce que Bramble a en tête est de dire qu'un sentiment distinctif désagréable est une propriété phénoménale des états mentaux telles que les expériences sensorielles. Je pense que nous pouvons comprendre ce sentiment de deux façons : i) en termes de sentiment unitaire, c'est-à-dire, c'est un même sentiment en vertu duquel toutes les expériences sensorielles désagréables sont qualifiées comme telles, et ii) le sentiment est accessible à l'introspection, c'est-à-dire, le sentiment distinctif fait partie de notre expérience consciente, au moins dans des circonstances normales.

Cependant, un des problèmes de la théorie est d'expliquer exactement comment le sentiment désagréable distinctif peut expliquer la force motivationnelle et normative des expériences hédoniques. Sobel (2005), par exemple, souligne le cas problématique des expériences agréables. Si des expériences agréables fournissent des raisons d'agir en vertu d'être phénoménales, il n'est pas évident de comprendre comment le fait d'être phénoménal explique pourquoi ces expériences nous donnent des raisons d'agir. Alston (1967) a une intuition similaire.

Je pense que la meilleure stratégie pour une théorie qui souhaite expliquer le caractère désagréable comme une propriété phénoménale est d'établir qu'une telle propriété phénoménale est mauvaise en soi. Autrement dit, d'établir que la différence entre le caractère désagréable et d'autres propriétés phénoménales consiste en ce que les expériences désagréables fassent mal. Comme le dit Smuts, « « être ressenti comme agréable » est aussi proche d'une expérience primitive que possible »[49] (Smuts, 2011, p. 11). Bramble n'a pas besoin d'expliquer les douleurs désagréables liées à la motivation et à la normativité en faisant appel à des croyances normatives, par exemple. Être désagréable est une propriété phénoménale intrinsèquement motivationnelle et normative car elle implique quelque chose de mauvais en soi. Ceci est donné. On n'a pas à expliquer pourquoi c'est comme ça, et il n'y a peut-être pas d'explication de cet état de fait.

3.2 Le problème de l'hétérogénéité pour la théorie du sentiment distinctif

L'hétérogénéité des expériences désagréables est un problème fondamental pour la théorie du sentiment distinctif. Selon le problème de l'hétérogénéité, même si toutes les expériences désagréables sont désagréables, il n'y a pas une propriété phénoménale unitaire, pas de sentiment unitaire, en vertu duquel toutes les expériences désagréables, et seulement elles, peuvent être qualifiées comme désagréables. La théorie du sentiment distinctif affirme le contraire, à savoir que toutes les expériences désagréables et elles seulement sont qualifiées comme désagréables parce qu'elles partagent le même et unique sentiment distinctif unitaire. Selon le problème de l'hétérogénéité, si on fait une liste de toutes les expériences qui sont désagréables et qu'on en fait l'introspection, on ne pourra pas trouver un seul sentiment distinctif de ces expériences qui explique ce en vertu de quoi ces expériences sont désagréables.

Bramble refuse une partie du problème de l'hétérogénéité : selon lui, même s'il semble qu'il n'y ait pas un sentiment unitaire parmi toutes les expériences désagréables et seulement elles, il existe en fait un tel sentiment. Cependant, ce sentiment n'est pas facile à identifier parce qu'il « imprègne » les expériences sensorielles. Ce caractère pénétrant du sentiment vise à expliquer pourquoi, après une introspection simple, il semble qu'il n'y ait pas une propriété phénoménale distincte qui unifie toutes les expériences sensorielles plaisantes ou désagréables, respectivement. En d'autres termes, le sentiment distinctif est là, c'est juste que nous ne pouvons pas en faire facilement l'introspection. Néanmoins, la

[49] Traduction de l'anglais.

notion de pénétrabilité, chez Bramble, pose problème. En bref, il est assez étrange d'affirmer que quelque chose est *distincti*f quoiqu'il passe *inaperçu*. Si quelque chose est distinctif, cela semble vouloir dire que ce quelque chose est tout du moins perceptible. C'est précisément l'un des avantages initiaux de la théorie du sentiment distinctif. Une introspection consciencieuse nous montre qu'il n'y a pas de sentiment désagréable unitaire partagé parmi toutes nos expériences désagréables.

La théorie du sentiment distinctif est utile pour illustrer un type d'approche non réductionniste du caractère désagréable comme propriété phénoménale. Nous pouvons expliquer que nous prenons un analgésique parce que le caractère désagréable de la douleur est mauvais en soi. Le caractère désagréable de la douleur étant mauvais, cela rend nos actions intelligibles, notamment car cela explique pourquoi une douleur désagréable est une raison motivante d'agir. Le caractère mauvais du caractère désagréable explique aussi pourquoi il est souhaitable de prendre un analgésique, pourquoi il est bon de faire cesser une expérience désagréable. Cependant, même si on accepte qu'une expérience avec cette propriété phénoménale est mauvaise en soi, la théorie du sentiment distinctif n'a pas réussi à rendre compte de l'hétérogénéité des expériences désagréables.

CHAPITRE 4 : LES THÉORIES DIMENSIONNELLES DU CARACTERE DÉSAGRÉABLE

Dans ce chapitre, j'expose d'autres théories qui partent de l'idée que le caractère déságréable est une propriété phénoménale. Ces théories, contrairement à la théorie du sentiment distinctif, ajoutent la notion de « dimension ». Selon une première théorie, *la théorie de la dimension hédonique*, le caractère désagréable est une dimension dans laquelle les expériences sensorielles peuvent varier. L'idée vient de Kagan (1992), même s'il utilise cette idée dans le cas des expériences agréables. Vu que la théorie de la dimension hédonique n'est pas capable de résoudre le problème de l'hétérogénéité, je considérerai une seconde théorie, qui donne une réponse à ce problème en utilisant la distinction entre déterminable et déterminant (Crisp, 2006). J'expliquerai la solution de Crisp au problème de l'hétérogénéité, qui se fonde sur la distinction déterminable-déterminant. Cependant, cette théorie rencontre elle aussi une difficulté fondamentale : elle ne peut pas rendre compte des cas de douleurs qui ne sont pas désagréables.

4.1 La théorie de la dimension hédonique

La théorie de la dimension hédonique propose que le caractère désagréable soit compris comme une dimension selon laquelle des expériences de douleur peuvent varier. Cette théorie repose sur une analogie : le volume du son est similaire à l'aspect agréable ou désagréable des états mentaux sensoriels (Kagan, 1992). La caractéristique principale de l'idée de Kagan est que si le volume est une dimension du son, alors le volume peut avoir des intensités différentes. Analogiquement, si le caractère désagréable est comme le volume, alors une expérience sensorielle peut avoir des intensités différentes.

Selon Bramble (2013), l'analogie de Kagan ne fonctionne pas. « Considérez que, pour la plupart des expériences, on puisse réduire complètement leur aspect agréable, tout en laissant l'expérience intacte, alors qu'on ne peut jamais réduire complètement le volume d'une expérience auditive et se retrouver avec l'expérience auditive en question"[50](Bramble, 2013, p. 209). Il y a cependant une façon simple et élégante de comprendre l'analogie de Kagan, en évitant cette critique de Bramble. Nous pourrions interpréter son analogie en comparant le caractère désagréable des *expériences*

[50] Traduction de l'anglais.

désagréables avec le volume du son. Si le volume est essentiel au son, nous pouvons affirmer qu'être désagréable est essentiel aux expériences mentales désagréables.

Il existe cependant un autre problème beaucoup plus important pour la théorie de la dimension hédonique : on ne peut pas rendre compte de l'hétérogénéité des expériences désagréables. Le problème avec la théorie de la dimension hédonique est que les variations du caractère désagréable ne peuvent pas être seulement expliquée en termes d'intensité. Alors que le volume du son varie seulement en intensité, le caractère désagréable des expériences ne varie pas seulement en intensité. La théorie de la dimension hédonique doit montrer que toutes les expériences désagréables et elles seulement sont unifiées en vertu d'être désagréables, en vertu d'avoir la même unité phénoménale, la même dimension phénoménale et que la seule variation parmi toutes ces expériences désagréables est en termes d'intensité. La théorie de la dimension hédonique n'est pas capable de répondre à cette intuition, et c'est pourquoi l'hétérogénéité des expériences désagréables est un problème pour cette théorie.

4.2 La solution au problème de l'hétérogénéité

Pour faire face à ce problème, tout en maintenant l'intuition que le caractère désagréable est quelque chose de ressenti, il a été proposé d'expliquer comment les expériences peuvent être agréables ou désagréables en faisant appel à la distinction déterminable-déterminant:

> Les expériences agréables diffèrent les unes des autres... Mais il y a une certaine qualité commune... elles sont vécues comme agréables... La distinction déterminable-déterminant nous aide aussi à clarifier le rôle du « vécu » dans cette analyse: être vécu comme agréable en tant que déterminable n'est pas un type particulier de sentiment vécu déterminé.[51] (Crisp, 2006, p. 109)

Je propose qu'une bonne manière de comprendre cette distinction c'est d'expliquer que les différentes déterminations d'un déterminable commun varient d'une façon non additive. Je propose que cette compréhension de la distinction déterminable-déterminant est très importante. Si être coloré est un déterminable, cela veut dire que ses différents déterminants, les propriétés de couleur, diffèrent les uns des autres de façon non additive :

[51] Traduction de l'anglais.

la différence entre être rouge et être bleu ne consiste pas à être coloré plus être *X*, pour être rouge, ou être coloré plus être *Y*, pour être bleu. Par contraste, considérons la propriété d'être un jeune animal. Tous les éléments qui appartiennent à cette catégorie ont un ingrédient distinctif, c'est-à-dire, être jeune, *plus* quelque chose d'autre. Par exemple, être un chaton signifie être jeune plus être un chat ; être un ourson, être jeune plus être un ours ; etc. Les propriétés d'être un chaton ou un ourson sont différentes façons d'être un jeune animal, et ils diffèrent les uns des autres d'une manière cumulative ; ils ne sont pas des déterminants d'un déterminable en commun.

Si nous essayons de comprendre l'idée de Crisp, on devrait établir que le caractère désagréable est un déterminable avec divers déterminants comme être une démangeaison, être une douleur, être une crampe, etc. C'est-à-dire qu'être une démangeaison et être une douleur ne partagent pas un ingrédient commun du désagréable plus quelque chose d'autre qui les rend différents. Ces sont des manières différentes d'être désagréables d'une façon non additive. Si nous comprenons le désagréable comme un déterminable, avec la compréhension non-additive que j'offre pour la distinction déterminable-déterminant, le problème de l'hétérogénéité est dissolu. Être une douleur ou être une démangeaison sont des façons d'être désagréables qui diffèrent de façon non additive : être une démangeaison et être une douleur ne partagent pas un sentiment désagréable unitaire. Nous pouvons expliquer qu'être désagréable est une propriété phénoménale qui est mauvaise en soi d'avoir, et aussi accepter qu'être désagréable n'est pas un ingrédient commun à toutes et seulement les expériences désagréables.[52]

Les dimensions sont les propriétés essentielles des déterminables et en vertu de ces dimensions chaque déterminant d'un déterminable en commun est différent du reste. Quand on parle de différentes manières d'être pour différents déterminants d'un déterminable commun, on fait référence aux variations parmi ces dimensions. Comparez cela aux différentes façons d'être un jeune animal. Les différences entre ces propriétés sont expliquées en termes de variations parmi *différentes* dimensions essentielles : seulement la propriété d'être un chaton implique être un chat, seulement la propriété d'être un ourson implique être un ours, etc. Si être une douleur et être une démangeaison sont des déterminants d'un déterminable commun, cela veut dire que ces propriétés partagent les mêmes dimensions essentielles.

---

[52] Pour plus sur la distinction sur determinables-déterminates voir Johnson (1921), Prior (1949), et Wilson (2009, 2017).

4.3 Les douleurs qui ne sont pas désagréables

Cette façon de comprendre la douleur par rapport au caractère désagréable implique que la douleur est nécessairement désagréable. Cependant, il ne semble pas nécessaire que la douleur soit désagréable. Les propriétés déterminées impliquent leurs déterminables. Si être écarlate est une propriété déterminée d'être rouge, être écarlate implique d'être rouge. Si le fait d'être une douleur est une propriété déterminée d'être désagréable, alors être une douleur implique d'être désagréable. Cela veut dire qu'être une douleur sans être désagréable est impossible, de la même manière qu'il est impossible d'être écarlate sans être rouge. Au contraire, il y a des cas qui suggèrent qu'il est possible d'éprouver des douleurs qui ne sont pas désagréables.

CHAPITRE 5 : EST-CE QUE LES DOULEURS SONT TOUJOURS DÉSAGRÉABLES ?

Dans ce chapitre, j'analyserai divers cas qui suggèrent que la douleur pourrait ne pas être désagréable. Si ces cas impliquent des expériences de douleur qui ne sont pas désagréables, cela montre que la douleur n'est pas nécessairement désagréable. Si être une douleur n'implique pas d'être désagréable, être une douleur ne peut pas être une déterminée du déterminable « être désagréable ». L'asymbolie à la douleur est l'exemple le plus clair où il semble que les gens ont des douleurs qui ne sont pas hédoniques, c'est-à-dire, des douleurs qui ne sont ni agréables ni désagréables. Cependant, je montrerai que l'idée qu'il existe des douleurs qui ne sont pas hédoniques peut être remise en question. Le fait que la douleur soit paradigmatiquement désagréable suggère que toutes les douleurs sont en fait toujours désagréables.

5.1 Douleurs non hédoniques

Certains philosophes ont émis l'idée que le caractère désagréable est une caractéristique nécessaire de la douleur: si ce n'est pas désagréable, ça ne peut pas être une douleur. Pitcher (1970), par exemple, a examiné divers cas qui remettent en question l'idée que la douleur soit nécessairement désagréable. Il soutient que ces cas pouvaient être expliqués par un certain modèle de la douleur - la théorie de la douleur de Melzack et Wall's (1965)[53] - et il a conclu qu'il ne trouvait aucun exemple de douleur qui n'était pas désagréable. Cependant, tout le monde n'était pas d'accord avec lui. Hall (1989), par exemple, a soutenu qu'on pouvait « avoir exactement les mêmes sensations que lorsqu'on est coupé, brûlé ou meurtri, et qu'elles ne seraient pas désagréables » (Hall, 1989, p. 643). Dans la suite, j'examinerai plusieurs cas qui suggèrent qu'on peut avoir des expériences de douleur non désagréables.

Il y a différents syndromes où il semble que les gens ne ressentent jamais de la douleur. Ces cas peuvent être considérés comme une *insensibilité congénitale à la douleur* (Cox et al., 2006; Juliao & Brotto, 1955; Sternbach, 1963). Il y a deux points clés concernant les cas d'insensibilité à la douleur congénitale. Tout d'abord, contrairement à ce que nous pourrions penser, certaines personnes diagnostiquées avec ce syndrome ressentent de la

[53] Voir la réimpression (Melzack & Wall, 1996).

douleur dans de rares occasions de leurs vies, et ces douleurs semblent être désagréables. Deuxièmement, il est possible que les patients souffrant d'insensibilité à la douleur congénitale aient des douleurs non désagréables. Cependant, cette éventualité n'est pas utile si nous cherchons des exemples concrets qui confirment que les gens peuvent avoir des douleurs non désagréables. Même si c'est possible, il n'existe aucune preuve concluante confirmant que les personnes souffrant d'insensibilité congénitale à la douleur aient des expériences douloureuses non désagréables.

D'autres exemples suggèrent que l'on peut perdre notre capacité à ressentir de la douleur, bien qu'étant nés avec cette capacité. Ce sont les cas *d'insensibilité acquise à la douleur*. Un exemple intéressant d'insensibilité acquise à la douleur est celui des soldats blessés qui ne semblent pas ressentir de douleur juste après avoir été sur le champ de bataille. Beecher (1956) a attiré l'attention sur ces cas grâce à son travail sur les soldats blessés à Anzio pendant la Seconde Guerre mondiale (par exemple : Bain, 2013; Klein, 2015b; Wall, 1999). Néanmoins, il n'y a aucune preuve de comportement ou de témoignage de soldats rapportant qu'ils ressentent de la douleur, mais que cette douleur n'est pas désagréable du tout. Même s'il est possible que les personnes ayant une insensibilité acquise à la douleur aient des douleurs qui ne soient pas désagréables, rien n'indique que ce soit effectivement le cas.

En outre, quelques personnes rapportent qu'ils ressentent de la douleur sans montrer la réaction typique: elles semblent être indifférentes à leur douleur. Par exemple, les cas de lobotomie, qui ont pour objectif de supprimer des douleurs insupportables, ont amené des philosophes à se demander si ces patients ont des douleurs non désagréables (e.g., Dennett, 1978; Pitcher, 1970). Les patients lobotomisés reconnaissent que leur expérience était similaire avant et après l'opération, mais ils cessent de manifester de la détresse et déclarent qu'ils ne sont plus dérangés par ces douleurs (Freeman & Watts, 1946; Watts & Freeman, 1948). Ils semblent devenus indifférents à leur douleur. Cette indifférence est expliquée par le fait que la douleur n'est plus désagréable. Néanmoins, je pense qu'une explication plus probable est que leurs douleurs sont encore désagréables. Par exemple, Trigg (1970) offre cette réponse pour expliquer les changements de comportement chez les patients lobotomisés par rapport aux leurs douleurs chroniques. Il souligne que les patients lobotomisés semblent avoir perdu la capacité de craindre et de s'inquiéter par rapport aux leurs douleurs. Étant donné que l'anxiété de vivre de forts épisodes douloureux semble avoir été la principale source de souffrance pour certains de ces patients, cela expliquerait pourquoi ils ont acquis de l'indifférence.

Si les patients lobotomisés perdent leur capacité à la crainte et au stress à la suite de leur lobotomie, ils peuvent maintenant être confrontés à leurs douleurs désagréables et être indifférents, ce qui n'implique pas que leurs douleurs ne sont pas désagréables. De plus, d'après l'explication de Trigg, Freeman et Watts affirment que: « La psychochirurgie modifie la réaction du sujet à la douleur sans changer matériellement sa capacité à ressentir de la douleur. La douleur peut être présente, mais elle est séparée de ses implications. »[54] (Freeman & Watts, 1946, p. 954–955). Ce qui explique leur changement de comportement, leur indifférence, n'est pas un changement par rapport à leurs expériences de la douleur, mais un changement dans leurs émotions à propos de ces douleurs qui restent désagréables.

Finalement, *l'asymbolie à la douleur* est une affection neurologique qui a attiré beaucoup d'attention dans la discussion philosophique sur la douleur (voir Bain, 2013; Corns, 2014; Grahek, 2007; Gray, 2014; Klein, 2015a; Vignemont, 2015). Les asymboliques rapportent que leurs douleurs ne leur font pas mal et parfois même ils sourient quand on leur applique des stimuli nocifs (Schilder & Stengel, 1928). Les asymboliques ne réagissent pas de la manière habituelle aux stimuli nocifs qui causeraient normalement de la douleur, même s'ils ressentent de la douleur. Cette pathologie est expliquée comme le résultat d'une lésion de l'insula postérieure, typiquement causée par des accidents vasculaires cérébraux et des tumeurs cérébrales à l'âge adulte (Berthier, Starkstein, & Leiguarda, 1988). Ce syndrome, pensent certains philosophes, fournit des preuves solides de l'existence de douleurs qui ne sont pas désagréables. Autrement dit, il pourrait y avoir des états mentaux qui instancient la propriété phénoménale d'être une douleur mais sans instancier aucune propriété phénoménale du caractère désagréable.

Plus précisément, quatre interprétations ont été données pour ce syndrome : i) celle de Grahek (2007), qui pense que les asymboliques ont des douleurs qui ne sont pas désagréables, ce qui explique leur comportement étrange; ii) celle de Klein (2015a), qui pense que les asymboliques ont encore des douleurs désagréables, mais qu'ils ont perdu leur capacité à se soucier de leur intégrité corporelle, ce qui explique leur comportement étrange; iii) celle de Bain (2013), qui pense que les douleurs des asymboliques ne sont pas désagréables, mais que les douleurs ne sont pas désagréables du fait que les asymboliques ont perdu leur capacité à se soucier de leur intégrité corporelle; et iv) celle de Gray (2014), qui pense que les asymboliques ont perdu leur capacité de ressentir de la douleur. Les solutions de Klein et de Bain semblent être disqualifiées en tant qu'explications possibles

[54] Traduction de l'anglais.

par les cas de somatoparaphrénie soulevés par De Vignemont (2015). Ces cas montrent que se soucier de son corps n'est pas nécessaire pour la motivation, contrairement à Klein, ni nécessaire pour que la douleur soit désagréable, contrairement à Bain. Nous devons donc décider entre l'interprétation de Grahek et celle de Gray.

Je pense qu'on doit opter pour la dernière pour une raison simple: la douleur est l'exemple paradigmatique d'une expérience désagréable. C'est-à-dire, il y a une intuition très forte que la douleur peut impliquer, au moins habituellement, le caractère désagréable. Si la douleur est paradigmatiquement un exemple d'une expérience désagréable, on doit avoir des preuves très convaincantes pour nous faire changer d'avis. Je pense que ce n'est pas le cas.

CHAPITRE 6 : MASOCHISME

Le masochisme est un cas intéressant pour comprendre la nature de la douleur et du caractère désagréable. C'est un cas intéressant car il semble impliquer deux choses contradictoires: i) la douleur est désagréable et nous devrions l'éviter à cause de ce caractère, et pourtant ii) les masochistes sont attirés par les expériences douloureuses, ils poursuivent ces expériences même si elles sont désagréables. L'existence de douleurs seulement agréables expliquerait pourquoi les masochistes sont attirés par ces expériences douloureuses et montrerait qu'il y a des douleurs qui ne sont pas désagréables. Plusieurs auteurs ont proposé différentes manières de comprendre ce qu'est le masochisme, mais je pense que ce qu'ils faisaient vraiment était décrire différents types de masochisme, différentes versions d'un phénomène commun. La caractéristique commune à tous ces types de masochisme est que les gens recherchent l'expérience de la douleur. Cependant, les raisons pour lesquelles les masochistes recherchent ces expériences peuvent varier. De plus, il ne semble pas que les gens recherchent des expériences douloureuses qui sont juste plaisantes. Il n'y a pas des preuves concluantes de l'existence de douleurs non désagréables dans la littérature sur le masochisme.

6.1 Masochisme moyen-but

La façon la plus courante dans la littérature philosophique d'expliquer le masochisme est de dire que les masochistes ne veulent pas vraiment avoir l'expérience de la douleur pour elle-même. En réalité, un masochiste est quelqu'un qui poursuit la douleur « comme un moyen de parvenir à quelque but... guidé par des raisons obscures comme la culpabilité ou le masochisme sexuel »[55] (Nagel, 1986, pp. 156–157).

Le mot « masochisme » a été introduit pour la première fois en médecine par Kraft-Ebing (1892). Le mot a été inspiré par le nom de l'écrivain Leopold Sacher-Masoch (1870).[56] D'après Kraft-Ebing, la caractéristique principale du masochisme n'était pas que les douleurs sont un moyen de gratification sexuelle, comme le masochisme est souvent

[55] Traduction de l'anglais.

[56] Voir la version réimprimée (Sacher-Masoch, 2004).

compris.[57] En fait, selon lui, les douleurs sont un moyen de se sentir soumis. Kraft-Ebing distinguait le masochisme du *sexual bondage*, ce dernier étant en fait beaucoup plus proche de ce que les définitions médicales actuelles comprennent par masochisme. Alors que dans le masochisme, tel que Kraft-Ebing le comprenait, l'expérience de la douleur semble avoir un rôle instrumental pour obtenir le sentiment de soumission, on en voit pas clairement comment la douleur a un rôle instrumental pour obtenir la satisfaction sexuelle dans le *sexual bondage*.

Il y a un autre exemple où les gens recherchent des douleurs et où ces expériences ont aussi un rôle instrumental. Goldstein (1983) défend l'idée que le masochisme se comprend mieux comme un exemple d'autopunition irrationnelle et anormale; le masochisme consisterait dans l'utilisation d'une douleur désagréable comme un moyen d'autopunition non méritée. Goldstein soutient que tous les cas d'autopunition impliquant de la douleur ne sont pas masochistes. Il pense que la différence entre un cas normal d'autopunition et le masochisme est que ce dernier est une autopunition irrationnelle. Dans le masochisme il n'y a pas de bonne raison d'être puni. Si, par exemple, vous vous punissez pour quelque chose que vous n'avez pas fait, cela compterait comme du masochisme d'après Goldstein.

Nous pouvons expliquer la *motivation* à s'engager dans une douleur désagréable, même si elle est désagréable: nous pourrions être motivés à poursuivre une douleur désagréable parce que nous croyons qu'une telle douleur désagréable apportera quelque chose d'autre que nous désirons. Nous pouvons également être *justifiés* à rechercher une douleur désagréable. Il peut y avoir des circonstances où il peut être souhaitable de s'engager dans une douleur désagréable comme un moyen. Je pense que le trait essentiel du masochisme moyen-but est que les gens poursuivent intentionnellement une douleur désagréable parce que cette expérience est désirée comme un moyen pour autre chose, même s'il n'y a pas de bonnes raisons de chercher ces douleurs désagréables comme moyen.

6.2 Douleurs d'effet collatéral

Il y a un autre type de cas qui pourrait facilement être confondu avec le masochisme moyen-but. Les gens peuvent supporter volontairement des douleurs désagréables, mais sans que l'expérience de douleur ne soit vraiment un moyen. Au contraire, la douleur est parfois un effet collatéral de ce qui est réellement poursuivi. Nous nous engageons parfois

[57] Deux bons exemples de ces approches contemporains du masochisme sont le manuel diagnostique et statistique des troubles mentaux (DSM-V) et la dixième révision de la classification statistique internationale des maladies et des problèmes de santé (ICD-10).

dans des activités qui impliquent des douleurs désagréables parce qu'elles sont un effet collatéral inévitable de quelque chose que nous poursuivons intentionnellement. Après une analyse plus minutieuse, nous pouvons se rendre compte que la douleur ne joue pas un rôle instrumental dans ces cas. Aller chez le dentiste est un bon exemple : on ne va pas chez le dentiste *parce que* ça fait mal, on y va *en dépit du fait que* ça fait mal.

Dans certaines occasions, nous supportons le caractère désagréable des expériences telles que la douleur parce que c'est l'effet collatéral d'un moyen ou d'un but. Nous pouvons être motivés à supporter un tel caractère désagréable car nous voulons le moyen ou le but, même si nous pensons que l'expérience désagréable est une conséquence inévitable de nos moyens ou de nos buts. En outre, il y a des occasions où nous pouvons avoir de bonnes raisons de supporter un tel caractère désagréable, et dans d'autres cas nous n'avons peut-être pas de bonnes raisons de le supporter, ce n'est pas souhaitable.

6.3 Masochisme de but en soi

Il y a différentes façons dont la douleur peut être un but en soi. Deux cas qui apparaissent dans la littérature philosophique sur le masochisme sont : i) une douleur désagréable qui fait partie d'un ensemble poursuivi, ou ii) une douleur désagréable qui est aussi agréable. Certains philosophes ont remarqué que, parfois, quelque chose de désagréable peut faire partie d'un tout qui est recherché. Pitcher (1970) offre un bon exemple du type de scénario qui serait considéré comme du *masochisme contextuel*:

> Imaginons que Agnew n'aime pas les olives, mais il aime les paellas qui les contiennent. Et nous n'imaginons pas qu'il tolère à peine la présence des olives détestées dans sa paella; de sorte qu'il aime le plat malgré eux: au contraire, nous devons supposer qu'il considère les olives comme constituant un ingrédient nécessaire pour une paella vraiment bonne — sans les olives, ce serait beaucoup moins intéressant.[58] (Pitcher, 1970, p. 484)

Cet exemple nous donne une idée de ce que cela peut signifier de rechercher une expérience douloureuse quand il s'agit de masochisme contextuel. C'est-à-dire, dans ce scénario, Agnew n'aime pas les olives en elles-mêmes, mais il les aime quand elles font partie de la paella, même si le goût des olives reste désagréable. Le goût des olives est désagréable pour Agnew quand il les mange, à l'écart des autres saveurs; il n'aime pas le

[58] Traduit de l'anglais.

goût parce que c'est désagréable. Cependant, quand Agnew mange des olives en paella, le goût de ces olives mérite d'être poursuivi lorsqu'il est accompagné d'une myriade d'autres arômes, textures, arômes, etc.

Le masochisme contextuel peut également se produire avec l'expérience produite en mangeant de la nourriture épicée. De la même manière que certains plats sont partiellement constitués de piments épicés, l'expérience résultant de la consommation de ces plats est en partie constituée par la sensation de brûlure dans la bouche lorsque nous mangeons le plat. La sensation de brûlure dans nos bouches que nous ressentons parfois lorsque nous mangeons de la nourriture épicée est causée par la capsaïcine contenue dans certains piments. L'épisode de dégustation d'un plat épicé implique une douleur brûlante désagréable comme l'une de ses parties essentielles et, en ces termes, telle expérience désagréable est poursuivie comme un but en soi.

Si le but dans son ensemble vaut vraiment la peine d'être poursuivi, c'est-à-dire, si le but est suffisamment bon, nous serons alors justifiés à nous engager dans quelque chose qui est partiellement composé d'une expérience désagréable. Si le tout est en effet assez bon, cela expliquera non seulement pourquoi nous pouvons nous engager dans ces expériences désagréables, mais cela constituera aussi une bonne raison de poursuivre le tout comme un but. Dans les cas de masochisme contextuel, qu'ils soient rationnels ou pas, les gens recherchent des douleurs désagréables qui sont des parties essentielles d'un tout. Cependant, le fait que la douleur ne soit pas recherchée comme un but, toute seule, mais comme une partie essentielle d'un but plus large, suggère que les douleurs impliquées dans le masochisme contextuel sont désagréables. Il n'y a pas aucune raison de douter que ces douleurs sont désagréables. Il existe différentes raisons de rechercher ces douleurs, mais aucune ne semble impliquer qu'elles soient seulement agréables.

Klein (2014) propose un autre type de scénario masochiste. Selon lui, certaines expériences telles que la douleur peuvent être à la fois agréables et désagréables. Klein appelle ces cas *plaisirs masochistes*. Klein pense que les plaisirs masochistes, y compris la douleur, sont agréables en plus d'être désagréables et parce qu'ils sont désagréables. Différentes raisons expliquent pourquoi le caractère désagréable de la douleur peut aussi être agréable dans ces situations. Le caractère désagréable d'une expérience douloureuse supportable peut devenir agréable: i) s'il est nouveau, ii) si cela nous permet d'exercer un contrôle de nous-même, iii) si cela représente une intimité particulière avec quelqu'un, iv) si cela repousse nos limites et nous aide à grandir et à changer, etc.

Les plaisirs masochistes éclairent une raison différente pour laquelle nous pouvons rechercher la douleur : nous pouvons rechercher une douleur désagréable si son caractère désagréable est aussi, en plus, agréable. C'est à cause de ce plaisir qu'on peut expliquer pourquoi nous pouvons être motivés et même justifiés à nous engager dans de tels plaisirs masochistes. Si le plaisir tiré d'expériences douloureuses est assez fort, nous pourrions être justifiés à poursuivre des plaisirs masochistes, même s'ils entraînent du déplaisir.

En conclusion, même s'il n'y a pas de cas concrets de douleurs non désagréables, pas même dans le masochisme, il semble au moins possible qu'il puisse y en avoir. Le fait que nous puissions concevoir la douleur comme n'étant pas hédonique suggère que nous devrions avoir une théorie du caractère désagréable qui permette d'expliquer cette possibilité.

CHAPITRE 7 : LA THÉORIE HÉDONIQUE DE DÉTERMINABLE-DÉTERMINÉE

Dans ce chapitre, je vais proposer une théorie qui a plusieurs vertus. Cette théorie maintient l'intuition que le caractère désagréable est quelque chose de ressenti, quelque chose de phénoménal; elle peut expliquer la possibilité de douleurs qui ne sont pas désagréables; elle peut rendre compte du cas du Ploner, un exemple où une expérience désagréable est dépourvu de son aspect phénoménal de douleur; elle peut expliquer comment deux expériences sensorielles peuvent varier de façon hédonique seulement, c'est-à-dire, l'une étant agréable et l'autre désagréable; enfin, cette théorie est capable de rendre compte de l'hétérogénéité des expériences désagréables. C'est la théorie hédonique de déterminable-déterminant (THDD).

Cependant, THDD doit faire face à deux difficultés. Premièrement, elle doit rendre compte de la variation phénoménale parmi les différentes façons d'être désagréable. THDD est capable de rendre compte de cet aspect. Deuxièmement, il n'est pas évident comment la propriété phénoménale d'être désagréable implique d'être mauvais en soi. Cependant, même si THDD n'explique pas pourquoi une expérience désagréable est mauvaise en soi, cette théorie est dans une meilleure position que les théories discutées dans les chapitres précédents.

7.1 La théorie hédonique de déterminable-déterminant

Voici ce que THDD propose pour rendre compte des expériences de douleurs désagréables :

> **Théorie hédonique de déterminable-déterminant du caractère désagréable de la douleur**
>
> Une expérience de douleur désagréable a deux propriétés phénoménales : i) la propriété phénoménale d'être une douleur, et ii) une propriété phénoménale déterminante (u1, u2, u3, etc.) de la propriété déterminable du caractère désagréable.

De cette façon, nous pouvons facilement rendre compte de la possibilité pour les asymboliques d'avoir des expériences douloureuses qui ne sont pas hédoniques. Autrement dit, ils auraient un état mental avec la propriété phénoménale d'être une

douleur, mais sans la propriété phénoménale d'être désagréable. Un autre avantage de dissocier la propriété d'être une douleur de la propriété d'être désagréable est que nous pouvons rendre compte des cas où, vraisemblablement, quelqu'un a une expérience désagréable, sans être une douleur ou une expérience phénoménale d'une autre forme sensorielle. Par exemple, certains chercheurs affirment que leurs « résultats démontrent, pour la première fois chez l'humain, une perte de la sensation de douleur avec un effet de douleur préservée »[59] (Ploner, Freund, & Schnitzler, 1999, p. 211). Si ce cas est réel, où même possible, THDD nous permet d'en rendre compte.

THDD a un autre avantage : nous pouvons expliquer comment différentes expériences sensorielles peuvent varier hédoniquement seulement. Nous pouvons expliquer comment le même type d'expérience sensorielle peut parfois être agréable et d'autres fois désagréable. Par exemple, nous pouvons avoir des expériences de goût de chocolat parfois agréables et parfois désagréables. Comment pouvons-nous expliquer cela ? L'expérience agréable de la dégustation de chocolat se compose de deux propriétés: i) la propriété goût-de-chocolat, et ii) une manière déterminée d'être agréable que j'appellerai chocolat-agréable. Ensuite, l'expérience désagréable de la dégustation de chocolat se compose de deux propriétés : i) la propriété goût-de-chocolat, et ii) une manière déterminée d'être désagréable que j'appellerai chocolat-désagréable. Deux expériences sensorielles différentes qui ne varient que hédoniquement ne sont pas des déterminants différents d'un déterminable commun. Les caractères agréable et désagréable de ces expériences sensorielles agréables et désagréables sont des propriétés déterminantes.

Finalement, nous devrions expliquer comment THDD résout l'hétérogénéité des expériences désagréables. D'après THDD, le caractère désagréable est phénoménal, mais il n'y a pas un sentiment unitaire du caractère désagréable car les différentes façons d'être désagréable varient de manière non additive. THDD est cohérente avec l'intuition qui soutient le problème de l'hétérogénéité. Plus précisément, toutes les expériences désagréables sont considérées comme telles en vertu de leur propriété déterminable d'être désagréable. Toutes les expériences désagréables sont désagréables d'une certaine manière, u1, u2, u3, etc. et chacune de ces manières d'être désagréable, chacun de ces caractères désagréables, varie par rapport aux autres d'une façon qui n'est pas additive. En d'autres termes, u1, u2, u3, etc., varient à travers les mêmes dimensions essentielles. Une fois que nous considérons qu'être désagréable est une propriété déterminable, nous pouvons expliquer comment les douleurs désagréables varient en fonction de leur

[59] Traduit de l'anglais.

caractère désagréable, même s'il n'y a pas un sentiment désagréable unitaire qui est partagé par toutes les douleurs désagréables.

7.2 Deux derniers problèmes pour THDD

Afin de donner une explication plus substantiel de THDD, nous devons clarifier comment nous pourrions savoir quelles sont les dimensions essentielles du caractère désagréable. Cela expliquerait comment des expériences désagréables partagent une propriété phénoménale commune, même s'il y a tellement de diversité concernant les différentes façons d'être désagréable. Comment pouvons-nous savoir quelles sont les dimensions des propriétés phénoménales? Pour répondre à cette question, nous pouvons faire appel à la notion d'espace qualitatif (voir Rosenthal, 2015).

En postulant un espace qualitatif, nous pourrions découvrir les dimensions essentielles des propriétés déterminables phénoménales. D'après Rosenthal (2015), nous pouvons utiliser la même méthodologie pour différentes modalités sensorielles. Par exemple, la même méthodologie a été utilisée pour essayer de créer un espace qualitatif pour les expériences olfactives (Young, Keller, & Rosenthal, 2014). Nous pourrions appliquer une stratégie similaire pour construire un espace qualitatif du caractère désagréable, au moins en principe. Même si nous ne connaissons pas exactement quelles sont les dimensions selon lesquelles le caractère désagréable peut varier, nous pouvons accepter que ces dimensions existent. « Nous pouvons certainement être en désaccord sur la dimension déterminante d'un type particulier, mais pourvu que ce désaccord soit raisonnable, l'existence même d'un tel désaccord permet de confirmer que nous partageons une compréhension intuitive du concept de dimension de détermination et de sa découverte. » (Funkhouser, 2014, p. 30)

Le dernier problème pour THDD est particulièrement problématique pour les théories qui prennent le caractère désagréable pour quelque chose de ressenti. Le problème est d'expliquer pourquoi le caractère désagréable est mauvais d'une manière non instrumentale et, comment, si le caractère désagréable est phénoménal, cette propriété phénoménale implique qu'une expérience désagréable soit mauvaise en soi. Je ne vais pas expliquer pourquoi le caractère désagréable implique un tel mal intrinsèque. Je le conçois comme un fait brut que le caractère désagréable implique quelque chose de mauvais non instrumentalement. Les expériences qui instancient le caractère désagréable ont une force motivationnelle et normative parce que avoir ces expériences sont mauvaises en soi. Alors, comment THDD fait-il mieux que les autres théories ?

Premièrement, THDD est préférable aux théories du contenu concernant la structure du caractère désagréable, c'est-à-dire, en ce qui concerne l'hétérogénéité de ce caractère. D'après les théories de contenu toutes les expériences désagréables sont désagréables de la même manière, comme si le caractère désagréable était un sentiment unitaire, comme si c'était un ingrédient que possèdent toutes les expériences désagréables et seulement elles, y compris les douleurs désagréables. Ce n'est pas le cas, comme le montre le problème de l'hétérogénéité. THDD est meilleure parce qu'elle rend compte de l'hétérogénéité des expériences désagréables.

Deuxièmement, THDD explique mieux la nature du caractère désagréable que les théories du contenu. L'un des principaux objectifs des théories du contenu, en faisant appel au contenu mental, était d'expliquer le fait que ce caractère est mauvais en soi. Cependant, on ne peut pas expliquer le caractère mauvais intrinsèque d'une douleur désagréable en fonction de quelque chose qui n'est pas mauvais, comme le montre le problème du meurtre du messager. Il est donc préférable de ne pas expliquer du tout ce caractère mauvais intrinsèque. THDD ne l'explique pas, mais n'entraîne donc pas d'incohérences.

Comparons maintenant THDD avec les théories du désir. Toutes sont capables de rendre compte de l'hétérogénéité des expériences désagréables. Alors, pourquoi devrions-nous préférer THDD? Je pense que nous devrions la préférer parce que cette théorie nous permet de garder l'intuition que le caractère désagréable est phénoménal, que c'est quelque chose de qualitatif, quelque chose de ressenti. Les théories du désir pourraient expliquer pourquoi il est instrumentalement bon d'avoir des sid-désirs en utilisant une explication évolutionniste et téléologique. Cependant, comme je l'ai expliqué au chapitre deux, il n'y a pas de raison motivante ou de raison non instrumentale de ne pas avoir de désirs pour des sensations de douleur hédoniquement neutres, car selon les théories du désir, une sensation neutre de la douleur n'a rien d'intrinsèquement mauvais. En revanche, THDD peut donner une explication assez simple à nos désirs concernant la douleur et son caractère désagréable. Selon THDD, nous sommes justifiés à désirer de ne pas avoir une douleur désagréable parce que le caractère désagréable est mauvais en soi, parce que les douleurs désagréables font mal. Le caractère désagréable d'après cette théorie est mauvais en soi et phénoménal.

Finalement, THDD se trouve dans une meilleure position que toutes les autres théories qui considèrent le caractère désagréable comme une propriété phénoménale sans donner une explication réductrice. En comparaison à la théorie du sentiment distinctif, du Chapitre 3, et de la théorie de la dimension hédonique, du Chapitre 4, THDD peut rendre compte de

l'hétérogénéité des expériences désagréables. La version développée à partir de la proposition de Crisp était également capable de gérer l'hétérogénéité des expériences désagréables. Cependant, THDD est dans une meilleure position concernant la relation entre la douleur et le caractère désagréable. THDD est également capable de rendre compte de: i) la possibilité de douleurs qui ne soient pas hédoniques, ii) des expériences désagréables qui ne sont pas sensorielles, comme le suggère le cas du Ploner, et iii) comment des expériences sensorielles peuvent varier seulement hédoniquement. Je conclus alors que cette approche est la meilleure option disponible pour comprendre la nature de la douleur et du caractère désagréable.

CONCLUSION

À partir des chapitres précédents, j'espère que le lecteur trouvera que THDD est la meilleure théorie disponible pour expliquer le caractère désagréable de la douleur. Selon cette approche, ce caractère est une propriété phénoménale qui implique d'être mauvais en soi. Les expériences douloureuses désagréables font mal, ce qui explique pourquoi nous sommes motivés et justifiés à éviter ces expériences. Le fait que les douleurs désagréables font mal nous donne une raison, et une bonne raison, de prendre un analgésique. Cette théorie nous permet d'expliquer la possibilité de douleurs qui ne sont pas désagréables, même si toutes les expériences douloureuses que nous avons semblent être désagréables. Cette approche nous permet également d'expliquer comment différentes expériences de diverses modalités sensorielles peuvent être agréables ou désagréables. Finalement, cette théorie permet de rendre compte de l'hétérogénéité du caractère désagréable.

BIBLIOGRAPHY

Alston, W. (1967). Pleasure. In P. Edwards (Ed.), *The Encyclopedia of philosophy* (Vol. 4, pp. 341–347). Macmillan.

Alvarez, M. (2016). Reasons for action: Justification, motivation, explanation. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016). Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/win2016/entries/reasons-just-vs-expl/

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders : DSM-5*. Arlington, VA: American Psychiatric Association.

Anscombe, G. E. M. (1957). *Intention*. Oxford: Blackwell.

Armstrong, D. M. (1962). *Bodily sensations*. Routledge.

Aydede, M. (2014). How to unify theories of sensory pleasure: An adverbialist proposal. *Review of Philosophy and Psychology*, *5*(1), 119–133.

Ayer, A. J. (1936). *Language, truth and logic*. London: V. Gollancz.

Bain, D. (2011). The imperative view of pain. *Journal of Consciousness Studies*, *18*(9–10), 164–185.

Bain, D. (2012). What makes pains unpleasant? *Philosophical Studies*, *166*(1), 69–89.

Bain, D. (2013). Pains that don't hurt. *Australasian Journal of Philosophy*, *92*(2), 305–320.

Bain, D. (2017). Why take painkillers? *Noûs*, (10.1111/nous.12228). https://doi.org/10.1111/nous.12228

Barber, J., & Bejenke, C. J. (Eds.). (1996). *Hypnosis and suggestion in the treatment of pain: a clinical guide*. New York: W.W. Norton.

Beecher, H. K. (1956). Relationship of the significance of wound to pain experienced. *Journal of the American Medical Association*, *161*(17), 1609–1613. https://doi.org/10.1001/jama.1956.02970170005002

Berthier, M., Starkstein, S., & Leiguarda, R. (1988). Asymbolia for pain: A sensory-limbic disconnection syndrome. *Annals of Neurology*, *24*(1), 41–49. https://doi.org/10.1002/ana.410240109

Boswell, P. (2016). Making sense of unpleasantness: Evaluationism and shooting the messenger. *Philosophical Studies*, *173*(11), 2969–2992.

Brady, M. (2017). Painfulness, desire, and the Euthyphro dilemma. *American Philosophical Quarterly*. Retrieved from http://eprints.gla.ac.uk/135993/

Bramble, B. (2013). The distinctive feeling theory of pleasure. *Philosophical Studies 162*, *2*, 201–217. https://doi.org/10.1007/s11098-011-9755-9

Corns, J. (2014). Unpleasantness, motivational *oomph*, and painfulness. *Mind and Language*, *29*(2), 238–254. https://doi.org/10.1111/mila.12048

Corns, J. (Ed.). (2017). *The Routledge Handbook of Philosophy of Pain* (1 edition). London ; New York: Routledge.

Cox, J. J., Reimann, F., Nicholas, A. K., Thornton, G., Roberts, E., Springell, K., … Woods, C. G. (2006). An SCN9A channelopathy causes congenital inability to experience pain. *Nature*, *444*(7121), 894–898. https://doi.org/10.1038/nature05413

Crisp, R. (2006). *Reasons and the good*. Oxford: Clarendon.

Cutter, B., & Tye, M. (2014). Pains and reasons: Why it is rational to kill the messenger. *Philosophical Quarterly*, *64*(256), 423–433.

Davidson, D. (1963). Actions, reasons, and causes. *Journal of Philosophy*, *60*(23), 685–700.

Dennett, D. C. (1978). Why you can't make a computer that feels pain. *Synthese*, *38*(3), 415–449.

Dretske, F. (1995). *Naturalizing the mind* (New Ed edition). Cambridge, MA: MIT Press.

Feldman, F. (2004). *Pleasure and the good life: Concerning the nature, varieties, and plausibility of hedonism*. Clarendon Press.

Fields, H. L. (1999). Pain: an unpleasant topic. *Pain*, *82*(Supplement 1), S61–S69. https://doi.org/10.1016/S0304-3959(99)00139-6

Finlay, S. (2007). Responding to normativity. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics, Vol. 2* (pp. 220–39). Clarendon Press.

Freeman, W., & Watts, J. W. (1946). Pain of organic disease relieved by prefrontal lobotomy. *The Lancet*, *247*(6409), 953–955. https://doi.org/10.1016/S0140-6736(46)91713-8

Funkhouser, E. (2006). The determinable-determinate relation. *Noûs*, *40*(3), 548–569.

Funkhouser, E. (2014). *The logical structure of kinds*. Oxford University Press.

Goldstein, I. (1983). Pain and masochism. *Journal of Value Inquiry*, *17*(3), 219–223.

Gracely, R. H. (1992). Affective dimensions of pain: How many and how measured? *APS Journal*, *1*(4), 243–247. https://doi.org/10.1016/1058-9139(92)90056-I

Grahek, N. (2007). *Feeling pain and being in pain*. MIT Press.

Gray, R. (2014). Pain, perception and the sensory modalities: Revisiting the intensive theory. *Review of Philosophy and Psychology*, *5*(1), 87–101. https://doi.org/10.1007/s13164-014-0177-4

Hall, R. J. (1989). Are pains necessarily unpleasant? *Philosophy and Phenomenological Research*, *49*(June), 643–59.

Hall, R. J. (2008). If It itches, scratch! *Australasian Journal of Philosophy*, *86*(4), 525–535.

Haybron, D. M. (2008). *The pursuit of unhappiness: The elusive psychology of well-being*. Oxford ; New York, NY: Oxford University Press.

Heathwood, C. (forthcoming). Unconscious pleasures and attitudinal theories of pleasure. *Utilitas*.

Heathwood, C. (2006). Desire satisfactionism and hedonism. *Philosophical Studies*, *128*(3), 539–563.

Heathwood, C. (2007). The reduction of sensory pleasure to desire. *Philosophical Studies*, *133*(1), 23–44.

Heathwood, C. (2011). Desire-based theories of reasons, pleasure, and welfare. *Oxford Studies in Metaethics*, *6*, 79–106.

Helm, B. W. (2001). *Emotional reason: Deliberation, motivation, and the nature of value*. Cambridge: Cambridge University Press.

Helm, B. W. (2002). Felt evaluations: A theory of pleasure and pain. *American Philosophical Quarterly*, *39*(1), 13–30.

Hill, H. E., Belleville, R. E., & Wikler, A. (1955). Studies on anxiety associated with anticipation of pain. II. Comparative effects of pentobarbital and morphine. *A.M.A. Archives of Neurology and Psychiatry*, *73*(6), 602–608.

IASP. (1986). Pain terms: A list with definitions and notes on pain. *Pain*, *3(Suppl)*, 216–21.

Jacobson, H. (2013). Killing the messenger: Representationalism and the painfulness of pain. *Philosophical Quarterly*, *63*(252), 509–519.

Johnson, W. E. (1921). *Logic*. Cambridge: Cambridge University Press.

Juliao, O. F., & Brotto, W. (1955). [Congenital generalized indifference to pain]. *Arquivos De Neuro-Psiquiatria*, *13*(4), 338–342.

Kagan, S. (1992). The limits of well-being. *Social Philosophy and Policy*, *9*(2), 169–189. https://doi.org/10.1017/S0265052500001461

Klein, C. (2007). An imperative theory of pain. *Journal of Philosophy*, *104*(10), 517–532.

Klein, C. (2012). Imperatives, phantom pains, and hallucination by presupposition. *Philosophical Psychology*, *25*(6), 917–928.

Klein, C. (2014). The penumbral theory of masochistic pleasure. *Review of Philosophy and Psychology*, *5*(1), 41–55.

Klein, C. (2015a). What pain asymbolia really shows. *Mind*, *124*(494), 493–516.

Klein, C. (2015b). *What the body commands: The imperative theory of pain* (1 edition). Cambridge, Massachusetts: The MIT Press.

Klein, C. (2017). Pain, care, and the body: A response to de Vignemont. *Australasian Journal of Philosophy*, *95*(3), 588–593. https://doi.org/10.1080/00048402.2016.1251478

Klein, C., & Martínez, M. (forthcoming). Imperativism and pain intensity. In D. Bain, M. Brady, & J. Corns (Eds.), *The Nature of Pain*.

Korsgaard, C. M. (1996). The origin of value and the scope of obligation. In O. O'Neill (Ed.), *The sources of normativity*. Cambridge: Cambridge University Press.

Krafft-Ebing, R. von. (1892). *Psychopathia sexualis: With especial reference to contrary sexual instinct: A medico-legal study*. Philadelphia, Pa. ; London: F.A. Davis.

Labukt, I. (2012). Hedonic tone and the heterogeneity of pleasure. *Utilitas*, *24*(02), 172–199.

Lycan, W. (2015). Representational theories of consciousness. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2015). Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/sum2015/entries/consciousness-representational/

Martínez, M. (2011). Imperative content and the painfulness of pain. *Phenomenology and the Cognitive Sciences*, *10*(1), 67–90.

Martínez, M. (2015). Pains as reasons. *Philosophical Studies*, *172*(9), 2261–2274.

Martínez, M., & Klein, C. (2016). Pain signals are predominantly imperative. *Biology and Philosophy*, *31*(2), 283–298.

Melzack, R., & Wall, P. D. (1996). Pain mechanisms: A new theory: A gate control system modulates sensory input from the skin before it evokes pain perception and response. *Pain Forum*, *5*(1), 3–11. https://doi.org/10.1016/S1082-3174(96)80062-6

Melzack, R., Wall, P. D., & Melzack, R. (1982). *The challenge of pain* (Rev. ed). Harmondsworth: Penguin.

Moore, G. E. (1903). *Principia ethica*. Cambridge: Cambridge University Press.

Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, *83*(October), 435–50.

Nagel, T. (1986). *The view from nowhere*. Oxford University Press.

Pitcher, G. (1970). The awfulness of pain. *Journal of Philosophy*, *67*(July), 481–491.

Ploner, M., Freund, H.-J., & Schnitzler, A. (1999). Pain affect without pain sensation in a patient with a postcentral lesion. *Pain*, *81*(1), 211–214. https://doi.org/10.1016/S0304-3959(99)00012-3

Price, D. D. (2000). Psychological and neural mechanisms of the affective dimension of pain. *Science*, *288*(5472), 1769–1772. https://doi.org/10.1126/science.288.5472.1769

Prior, A. N. (1949). Determinables, determinates and determinants. *Mind*, *58*(229), 1–20.

Rainville, P., Carrier, B., Hofbauer, R. K., Bushnell, M. C., & Duncan, G. H. (1999). Dissociation of sensory and affective dimensions of pain using hypnotic modulation. *Pain*, *82*(2), 159–171. https://doi.org/10.1016/S0304-3959(99)00048-2

Ramachandran, V. S., & Hirstein, W. (1998). The perception of phantom limbs. The D. O. Hebb lecture. *Brain*, *121*(9), 1603–1630. https://doi.org/10.1093/brain/121.9.1603

Rosenthal, D. M. (1986). Two concepts of consciousness. *Philosophical Studies*, *49*(May), 329–59.

Rosenthal, D. M. (2015). Quality spaces and sensory modalities. In S. Coleman (Ed.), *Phenomenal qualities: sense, perception, and consciousness* (First edition). Oxford: Oxford University Press.

Ruben, D.-H. (2009). Con-reasons as causes. In C. Sandis (Ed.), *New Essays on the Explanation of Action* (pp. 62–74). Palgrave-Macmillan.

Sacher-Masoch, L. R. von. (2004). *Venus in furs*. Barnes & Noble Publishing.

Schilder, P., & Stengel, E. (1928). Schmerzasymbolie. *Zeitschrift Für Die Gesamte*, *113*(1), 143–158.

Schroeder, T. (2004). *Three faces of desire*. Oxford: Oxford University Press.

Schroeder, T. (2017). Desire. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2017). Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/sum2017/entries/desire/

Searle, J. R. (1979). *Expression and meaning: Studies in the theory of speech acts*. Cambridge: Cambridge University Press.

Siegel, S. (2016). The contents of perception. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016). Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/win2016/entries/perception-contents/

Smith, M. (1995). *The moral problem*. Malden, Mass. ; Oxford: Blackwell.

Smuts, A. (2011). The feels good theory of pleasure. *Philosophical Studies*, *155*(2), 241–265.

Snyder, S. H. (1996). *Drugs and the brain* (New ed.). New York: Scientific American Books.

Sobel, D. (2005). Pain for objectivists: The case of matters of mere taste. *Ethical Theory and Moral Practice*, *8*(4), 437–457. https://doi.org/10.1007/s10677-005-8839-z

Sternbach, R. A. (1963). Congenital insensitivity to pain: A critique. *Psychological Bulletin*, *60*(3), 252–264. https://doi.org/10.1037/h0042959

Trigg, R. (1970). *Pain and emotion*. Oxford: Clarendon P.

Tye, M. (1995). *Ten problems of consciousness: A representational theoryof the phenomenal mind*. Cambridge, Mass. ; London: MIT Press.

Tye, M. (2006). Another look at representationalism about pain. In M. Aydede (Ed.), *Pain: New Essays on Its Nature and the Methodology of Its Study*. Cambridge, Mass: MIT Press.

Vignemont, F. de. (2015). Pain and bodily care: Whose body matters? *Australasian Journal of Philosophy*, *93*(3), 542–560.

Wall, P. D. (1999). *Pain: The science of suffering*. London: Weidenfeld & Nicolson.

Watts, J. W., & Freeman, W. (1948). Frontal lobotomy in the treatment of unbearable pain. *Research Publications - Association for Research in Nervous and Mental Disease*, *27 (1 vol.)*, 715.

Wilson, J. (2017). Determinables and determinates. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2017). Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/spr2017/entries/determinate-determinables/

Wilson, J. M. (2009). Determination, realization and mental causation. *Philosophical Studies*, *145*(1), 149–169.

Wood, J. N. (Ed.). (1993). *Capsaicin in the study of pain*. London: Academic Press.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization.

Young, B. D., Keller, A., & Rosenthal, D. (2014). Quality-space theory in olfaction. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00001

Le Caractère Désagréable de la Douleur

Résumé

La douleur est désagréable. Étant donné que la douleur est l'exemple paradigmatique d'une expérience désagréable, mon but est de clarifier ce qui caractérise la douleur et le caractère désagréable. J'essaie notamment de comprendre ce que peut signifier qu'une douleur soit désagréable et ainsi d'élucider la structure des expériences désagréables. Ce faisant, j'aborde plusieurs aspects problématiques de la relation entre la douleur et le caractère désagréable des expériences. Je fournis également une compréhension générale de ce que signifie pour une expérience non nécessairement douloureuse d'être désagréable.

Cette thèse s'organise en sept chapitres correspondant à trois problématiques principales : i) qu'est-ce qui constitue le caractère désagréable de la douleur ? (Chapitres 1 & 2), ii) comment rendre compte de la grande diversité phénoménale du caractère désagréable des expériences ? (Chapitres 3 & 4), et iii) dans quels cas la douleur n'est-elle pas désagréable ? (Chapitres 5 & 6). Dans le dernier chapitre (Chapitre 7), j'offre une réponse générale aux trois problématiques principales en proposant ma propre théorie sur le caractère désagréable de la douleur. D'après cette théorie, une expérience désagréable est une expérience ressentie, le caractère désagréable est défini comme une propriété phénoménale des états mentaux, et cette propriété doit être comprise en utilisant la distinction déterminable-déterminant. Mon travail fournit ainsi une compréhension détaillée de la nature de la douleur et du caractère désagréable.

Mots-clés : douluer ; désagréable ; hétérogénéité ; contenu ; désir ; déterminable ; déterminant ; masochisme

The Unpleasantess of Pain

Summary

Pain is unpleasant. Given that pain is the paradigmatic example of an unpleasant experience, I aim to shed light on what pain and unpleasantness are by trying to understand what it means for a pain to be unpleasant, what the structure of unpleasantness is, and by tackling several problematic aspects of the relation between pain and unpleasantness. By doing this, I will also provide a general account of what it means for an experience that might not be a pain to be unpleasant.

The thesis is organised into seven chapters and divided by three main themes: i) what the unpleasantness of pain consists in, ii) how we can account for the great phenomenal diversity among experiences of unpleasantness, and iii) which cases suggest that there could be pains that are not unpleasant. Broadly, the first two chapters deal with the first theme, the third and fourth chapter with the second theme, and the fifth and sixth chapter focus on the third theme. In the final chapter I offer a conclusion of the three main themes by providing my own view on the unpleasantness of pain. According to this account, an unpleasant experience is something felt, it is a phenomenal property of mental states, and this property should be understood using the *determinable-determinate* distinction. By doing all of this, this thesis will provide a detailed understanding of the nature of pain and unpleasantness.

Keywords: pain; unpleasant; heterogeneity; content; desire; determinable; determinate; masochism

UNIVERSITÉ PARIS-SORBONNE
ÉCOLE DOCTORALE :
ED 5 – Concepts et langages
Maison de la Recherche, 28 rue Serpente, 75006 Paris, FRANCE
DISCIPLINE : Philosophie