

UNIVERSITY OF GRONINGEN
FACULTY OF PHILOSOPHY

WITTGENSTEINIAN ARGUMENTS
AGAINST A CAUSAL THEORY
OF REPRESENTATION

Master's Thesis ResearchMaster KKD

Author: Jan Willem Wennekes

Supervisors: Prof. dr. M.R.M. ter Hark, prof. dr. M.J.B. Stokhof

Third Reader: Dr. F.A. Keijzer

Date: August, 2006

Abstract

Daniel Dennett and Fred Dretske are two well known philosophers who have written substantially on representation (also referred to as intentionality or aboutness). Both are under the impression that the correct solution for the problem of representation (i.e. how do representations manage to represent something?) will have to be constituted by a causal explanation of the processes at work. In their philosophy, they have embraced the hypothetico-deductive method of the natural sciences, which functions against the background of a causalist view on the world. Both Dennett and Dretske describe simple 'intentional systems', such as thermostats as having an internal state that is about an external world property (in this case the natural magnitude temperature) because of some essential causal relation between the internal state and the outside world property represented by it. I therefore characterize their views as the causal theory of representation, since they try to find a causal explanation for representation.

In the 1930's, Ludwig Wittgenstein already criticized Russell and Ogden and Richards for similar viewpoints (although there are some dissimilarities). The aim of my thesis is to apply and extend this criticism to the contemporary theories of Dennett and Dretske. In my thesis I first describe their viewpoints on representation. I then give an overview of Wittgenstein's method of philosophy, which constitutes an entirely different approach from the scientific method both Dennett and Dretske have embraced. Instead of approaching the problem from a scientific point of view, Wittgenstein believes that the 'problem' of representation originates in a misunderstanding of our own language, e.g. by assuming that "because my muscles contracted" and "because I felt like it" are explanations of the same type to the question "why did you hit him?". Two of my main arguments against the causal theory of representation are that it fails to make the distinction between (1) causes and reasons and (2) internal and external relations. Failure to appreciate these distinctions results in a number of problems within the causal theory of representation. In the fourth chapter of my thesis I discuss these differences and show how they throw a different light on the nature of representation.

Acknowledgements

When I finished Artificial Intelligence at the University of Groningen, there were still some philosophical issues left unsolved for me. Thus, while working as a consultant, I decided to follow up on these issues: I quit my job and became a full-time student again. The one thing driving me back to philosophy, as it were, was Wittgenstein's work. So naturally, when the time came to choose a subject for my final thesis, the choice was in that direction. Analysing problems and trying to find that particular understanding of it that makes them almost silly and simple on the one hand, but ever so complicated on the other, has always been a very attractive aspect of philosophy to me. And Wittgenstein displays brilliant examples of that approach.

This thesis would not have come to the point where it is right now without the help and advice I received from two Dutch Wittgenstein scholars: Michel ter Hark and Martin Stokhof. I wish to thank them for their patience, helpful suggestions and good advice. Thank you for your time and commitment to this project. I also wish to thank the third reader of this thesis, Fred Keijzer, for lending me in his time and his knowledge of both Dennett and Dretske.

I am also grateful to other people for their criticism on earlier drafts, their advice in discussions and interest in the project as a whole. Especially helpful were my conversations with and the advice received from Kim van Gennip, Marie McGinn, and Jeanne Peijnenburg. Thank you all for your support.

Last but not least, a big thank you to my friends and family who have helped and supported me throughout the writing of this thesis.

University of Groningen
Groningen, August 2006

Table of Contents

CHAPTER 1: INTRODUCTION.....	9
1.1 MOTIVATION	9
1.2 RESEARCH QUESTION	11
1.3 RESEARCH METHOD	12
1.4 OUTLINE	12
CHAPTER 2: THE CAUSAL THEORY OF REPRESENTATION	17
2.1 INTRODUCTION	17
2.1.1 Chapter Outline.....	18
2.2 THE PROBLEM OF REPRESENTATION	19
2.3 THE CAUSAL WORLD PICTURE.....	22
2.4 THE CAUSAL THEORY OF REPRESENTATION.....	24
2.4.1 Representation and Intentionality.....	25
2.4.2 Dennett	27
2.4.3 Dretske	31
CHAPTER 3: WITTGENSTEIN'S METHOD	39
3.1 INTRODUCTION	39
3.2 WITTGENSTEIN'S STYLE.....	40
3.3 THE NEGATIVE ROLE OF PHILOSOPHY	41
3.3.1 The source of philosophical problems: the theoretical attitude.....	42
3.3.2 Misled by the form of our language: the character of depth	44
3.4 THE POSITIVE ROLE OF PHILOSOPHY	46
3.4.1 Ploughing through the whole of language	46
3.5 THERAPEUTIC APPROACH.....	48

3.6	METHODS AND TECHNIQUES	49
3.6.1	Family resemblance	49
3.6.2	Language games and form of life	51
3.6.3	Grammar	53
CHAPTER 4: A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION		57
4.1	INTRODUCTION	57
4.1.1	Chapter Outline	58
4.2	METHODOLOGY	60
4.2.1	Philosophy deals with a different category of problems	60
4.3	CAUSES VERSUS REASONS	63
4.3.1	Dennett and Dretske on causes and reasons	63
4.3.2	The difference between causes and reasons	65
4.3.3	Understanding is not a form of interpretation	71
4.4	INTERNAL VERSUS EXTERNAL RELATIONS	75
4.4.1	Science and Philosophy revisited	75
4.4.2	Problems with the indication approach	80
4.5	A POSITIVE OUTLOOK	84
4.5.1	What it is to understand a concept	84
4.5.2	Learning a language is learning a form of life	86
CHAPTER 5: CONCLUSIONS		93
5.1	RESEARCH OUTCOME	93
5.2	SUGGESTIONS FOR FUTURE RESEARCH	97
BIBLIOGRAPHY		101
ABBREVIATIONS		101
FURTHER REFERENCES		101

"We can't solve problems by using the same kind of thinking we used when we created them."

- Albert Einstein

Chapter 1

Introduction

1.1 Motivation

“I should have liked to produce a good book. This has not come about, but the time is past in which I could improve it.” (PI, preface)

These are the final words of the preface Wittgenstein wrote to his *Philosophical Investigations*, one of the most influential books in twentieth century philosophy. To have written something that displays such a deep insight into the nature of philosophical problems, the forms of our language and the way our mind functions, yet at the same time to feel this humble or even doubtful about it, is typical for Wittgenstein. It shows how deeply engaged Wittgenstein was with philosophy. Many other works, notes and commentaries were to follow this book. One reason why Wittgenstein is intriguing as a philosopher is that he has something to say about virtually any philosophical subject. Avrum Stroll, who has written substantially on Wittgenstein, describes him as follows:

More than any other analytical philosopher, he has changed the thinking of a whole generation. Like Plato, Aristotle and Kant, he is the product of an enormous subliteration of commentaries: by Max Black, Garth Hallet, Eike von Savigny, Gordon Baker, and P.M.S. Hacker, among others. The quantity of his work, unlike the tiny amount of material produced by Frege and Austin, is enormous and comparable in size to that of Plato and Aristotle. Moreover, it covers the entire gamut of philosophy, from logic through philosophical psychology to considerations of culture and value. It has also had a profound effect on the nonphilosophical disciplines in the humanities and social sciences. (Stroll, 2000, p. 252-523).

CHAPTER 1 – INTRODUCTION

Indeed, since first reading the book it has never loosened its grip on my mind and I have not been able to think coherently in a different way about philosophical issues from that point on. It had to be Wittgensteinian¹. Consequently, there remained no other option for me than to write my final thesis with the aid of Wittgenstein's method and remarks. I can only hope that the reader will appreciate my passion for Wittgenstein's philosophy and that it will be reflected in my writings appropriately.

Some of the material in this thesis can be related to earlier work of mine on Wittgenstein in various essays, written mainly in the past 3 years. I've worked on Wittgenstein – I would almost like to say, with Wittgenstein – various times and each time his insights into the problems at issue are profound and surprisingly deliberating. It is as if, for each philosophical problem that seems to imprison you like a deep, dark pit, Wittgenstein has developed a way – a ladder so to say – to get out of the darkness. The issue of how language connects up with the world – or the issue of representation as I will call it in this thesis – is one of the most pervasive and deeply rooted issues that surfaces again and again in Wittgenstein's work. Language has been a subject for Wittgenstein throughout his entire life and philosophical writings. I've now come to believe that Wittgenstein already had the substantial insights known from the *Philosophical Investigations* at the time of writing the *Tractatus*, even though most commentators attribute this to the period leading to his *Philosophical Investigations* and thereafter². Notwithstanding, there are a number of differences between these both works but they are, in my opinion, not as substantial as is commonly thought. Still, even if a number of commentators will not agree that the *Tractatus* defends a likewise view on how language and the world are related, it remains obvious that language was one of Wittgenstein's main subjects in that first book. And on consideration of other works by Wittgenstein, such as *On Certainty*, we see that the issue resurfaces in different ways again and again. Wittgenstein's insights into the workings of language are his main tool to analyse and deconstruct philosophical problems. These considerations have led me to settle on the issue of *representation* as the main subject of my Master's thesis. I have now had the opportunity to combine my previous knowledge and insights into Wittgenstein's viewpoints into one product. If I have succeeded in representing Wittgenstein's profound insights faithfully is something that remains for the reader to judge.

¹ To outsiders this stubbornness of Wittgensteinian thinkers is often frustrating. However, I am not alone in this: "Serious students of Wittgenstein will never approach philosophical issues in pre-Wittgensteinian ways" (Stroll, 2000, p. 254).

² See for an excellent and convincing plead for this reading of the *Tractatus* Marie McGinn's forthcoming book "Elucidation in Wittgenstein's *Tractatus*", Oxford University Press.

1.2 Research Question

This thesis is based on a Wittgensteinian method of doing philosophy. Because this method is at odds with the traditional outlook of philosophers, this thesis might in some respects differ from a traditional philosophical thesis. The thesis started from a philosophical interest in the nature of representation and current philosophical debates. The fact that current philosophers still make the type of conceptual mistakes Wittgenstein was opposing in his time, made that I wanted to grasp the opportunity to show the errors still present in these contemporary views on representation. I have attempted to apply Wittgenstein's method and remarks to the best of my knowledge, with this thesis as the result. The following general research question reflects the investigation within this thesis:

“What is the nature of representation?”

Here, “representation” has to be taken as analogous with “meaning” or “intentionality” (this will be explained fully in Chapter 2). This research question obviously has a long tradition within philosophy and attempts to solve it have been made numerous times, in many different ways. It can be rephrased for example as “how does language relate to the world?” or “how can anything in the mind be a representation of anything outside the mind?”³. These questions have both led me towards the contemporary viewpoints on representation as well as (back) to Wittgenstein's views. It is, so to say, the common thread that runs throughout this whole thesis, and possibly, throughout my entire philosophical development. As general as this research question is, it can be specified by rewriting it as follows for the specific aim of this thesis:

“What critique can be given on contemporary philosophical theories on representation using Wittgenstein's remarks?”

Besides answering this research question, I also wanted to give an overview of Wittgenstein's method of philosophy, as I think it deserves the consideration of contemporary philosophers.

³ Cf. Putnam (1992) pp. 21-22 where he discusses e.g. philosophers in general and Kant in particular asking a similar question. Also see Crane (1995) who devoted his first chapter to such questions.

1.3 Research Method

The above stated research question and my earlier remarks on motivation both show that I am committed to Wittgenstein's method of philosophy. Wittgenstein himself calls this method a grammatical investigation, but this should not be taken as an indication that the investigation is purely a linguistic one. Wittgenstein is indeed known as a linguistic philosopher, but there is more to his philosophy than that. One could question whether the label "linguistic philosopher" in fact applies to Wittgenstein as much as it is used. It is obvious that Wittgenstein often concentrates on language, its use, concepts and their mutual relations, but his aim is to use these in order to show the actual issues he is concerned with. So he will investigate the language related to "length" if he is concerned with that *length* itself actually is. Yet this does not amount to a metaphysical investigation. For now these remarks will have to suffice as we will look into this subject later on.

Those convinced of Wittgenstein's method and philosophy will agree that this method is the technique *par excellence* for my aim, others philosophers might disagree. It sometimes seems as if this amounts to an almost religious matter as to what is the correct way of doing philosophy. Unfortunately I lack the time and space to go into this subject in this thesis, however interesting it is, both to philosophy in general as to the reductionistic programme currently pursued by many philosophers. For now I will have to restrict myself and simply state that Wittgenstein's philosophical method is what I see as the most fitting and fruitful way to investigate the phenomenon of representation. Finally I wish to remark that readers unfamiliar with Wittgenstein's method and views on philosophy are directed to Chapter 3, where I will give an outline of these two related issues.

1.4 Outline

In the next chapter I will provide an outline of what I have called the *causal theory of representation*. In philosophy, the problem of representation is a widespread and ancient problem, which has been discussed numerous times and is still the subject of contemporary debate. Connected to this is the question of how mental states such as beliefs, thoughts and desires are able to represent (as they are also seen as representations). Questions posed take the form of "How is it possible that the word "red" means red?" and "How can a thought represent something other than itself, something outside of the head?". Repeatedly I have come into contact with works by analytical philosophers, such as Dennett and Dretske, who try to deal with these questions by offering an explanation for the mechanisms responsible for the phenomenon. In the

CHAPTER 1 – INTRODUCTION

chapter I will discuss the solution offered by these contemporary philosophers and give an outline of their respective viewpoints. Both philosophers have developed their views against the background of the causal world picture, meaning that they believe that representation is in fact a causal process and is thus in need of a causal explanation. This outline will serve as the basis for the critique I will offer in Chapter 4.

Chapter 3 will provide an overview of Wittgenstein's method of philosophy and his rationale for this unusual way of dealing with philosophical problems. An important issue will be the difference Wittgenstein sees between the methods of the (natural) sciences and philosophy, a difference that also influences the domain of phenomena of these two disciplines: whereas scientific problems can be solved by formulating hypothesis and experimentation, philosophical problems are of a whole different nature: they are conceptual rather than empirical which means that no experiments can help to solve the problem. Rather, the problem has come into existence because we have misunderstood the forms of our language. I will focus on two roles Wittgenstein's philosophy has. The first role has a more negative character as it concerns the criticism his philosophy offers towards philosophical misconceptions. The second role, however, is of a more positive nature: besides offering criticism Wittgenstein does in fact show us a way of looking at certain phenomena that delivers an insight into the true nature of these phenomena. Furthermore I will discuss the therapeutic nature of Wittgenstein's remarks. This aspect of his writing is related to his style of philosophy: instead of explicitly presenting a clear theory to the reader as his solution to a problem, Wittgenstein regularly invites the reader to follow in his footsteps and consider numerous different but sometimes overlapping questions, thought-experiments or everyday examples. Finally I will give a short overview of some important methods and concepts within Wittgenstein's work.

After having discussed both the causal theory of representation and Wittgenstein's philosophical method I will bring the two together in Chapter 4. This chapter will provide my criticism of the causal theory of representation, based on Wittgenstein's remarks. First off, I will discuss the differences between the method followed by the proponents of the causal theory of representation (i.e. the method followed by both Dennett and Dretske) and the method followed by Wittgenstein. They are fundamentally different, which is reflected in the type of solution to the problem of representation offered by both sides of the debate. Subsequently I will discuss the misconceptions that can be located within the causal theory of representation with the help of Wittgenstein's philosophy. My main arguments will be that Dennett and Dretske do not distinguish properly between causes and reasons on the one hand, and internal and external relations on the other. This results in a number of conceptual problems for the causal theory of representation. These points can be seen as the negative contribution of Wittgenstein's philosophy to the problem of representation as they form the criticism to the view supported by Dennett and Dretske. Accordingly, I will also concentrate on the positive contribution

CHAPTER 1 – INTRODUCTION

Wittgenstein's philosophy has to offer to the problem of representation. As a conclusion to this chapter I will thus focus on aspects such as training, learning and the use of concepts in everyday practices and internal relations.

Finally, Chapter 5 will provide the reader with a short overview of this thesis, summarize my conclusions and provide recommendations for further research.

“He who shall duly consider these matters will find that there is a certain bewitchery or fascination in words, which makes them operate with a force beyond what we can naturally give account of.”

- Robert South

Chapter 2

The Causal Theory of Representation

2.1 Introduction

Philosophers have long wondered about the nature of the relation between language and reality. How is it possible, we ask, that a word *means* points to something outside itself? Or for that matter, how is it possible that a drawing, picture, photo or diagram points to something outside itself? The same questions have been asked about the relation between thought and reality. How is it possible that a thought is *about* something? Both types of questions fall under the problem of intentionality, or representation. For example, how is it possible that the word ‘apple’ means an apple, or that a drawing of an apple represents an apple? How is it possible that my thought is about an apple, e.g. when I desire an apple? That is, how can one thing stand in for another thing?⁴

A widespread and at the outset intuitively appealing explanation is offered by referentialism. Naïve forms of referentialism hold that the meaning or content of a word – or of any other representation – is its standing in for a certain object. It is this picture of how language functions that Wittgenstein criticizes in the opening remarks of the *Philosophical Investigations*. In PI 1, after describing Augustine’s picture of language, he writes: “In this picture of language we find the roots of the following idea: Every word has a meaning. This meaning is correlated with the word. It is the object for which the word stands.” However, this conception of meaning has a fundamental defect: false representations. This problematic issue can be traced back at least to Democritus and

⁴ Note that saying that a representation is about something other than itself (e.g. that a word is about something other than itself) and that this is mysterious, as Dennett and Dretske seem to be putting forward, is actually a very queer phrase! What is it supposed to mean? Are representations ever about themselves? If you say that X is “about Y” logic implies that it is also possible for X to be “about Z”, so if you say that it is strange for X to be “about something else” you are logically implying that X is generally only “about X”. Otherwise, why would you need to add this information?

CHAPTER 2 – THE CAUSAL THEORY OF REPRESENTATION

Plato (cf. Glock, 1996, p. 184). It was Plato who formulated this problem concerned with representation in the *Theaeteteus*, when he asked how it is possible to think or state something that is not the case. For if something is not the case, e.g. if we say “there is a unicorn in the garden”, how can one refer to it? Since there is no unicorn in the garden, what does the corresponding thought or sentence relate to? What object does it stand for?

These issues form the problem of representation, which I will further introduce in this chapter. I will also discuss a theory, which I have named ‘the causal theory of representation’, that has been proposed as a solution to this problem. This theory will be the subject of my critique in Chapter 4. As representatives of the causal theory of representation I will discuss both Fred Dretske and Daniel C. Dennett. The causal theory of representation tries to solve the problem of representation by explaining the nature of the relation between a representation and that which it represents in terms of causal processes. The term ‘causal’ is part of the conception of meaning as something that can and must be fitted in the causal world picture. This view, which is roughly the idea that all phenomena can be explained in terms of causes and their effects, will also be discussed in this chapter. Combined with the problem of representation, it results in ascribing causal powers to representations themselves.

What is important to note is that in this chapter I will describe both the problem of representation and the (alleged) solution to it, within the framework of the causal world picture. The philosophers I wish to discuss hold the latter to be *the* framework that will provide the necessary background to their research. In describing the problem, we will thus have to adopt the attitude towards representation, language and thought that fits within this causal world picture. It is within this idea of the world that the question of how a purely physical system, such as the brain, is able to represent things has arisen. This question is what the causal theory of representation claims to solve. The point is however, as I hope to show in chapter 4, that this approach towards the problem of representation is itself part of the reason why representation seems to be so problematic in the first place. It is in fact part of the problem and needs to be scrutinized as much as the causal theory of representation. But before doing so, we will first take a close look at the theory itself.

2.1.1 Chapter Outline

I will start by explaining the problem of representation, the main subject of this thesis: how is it possible that one thing represents another? I will then discuss the causal world picture, a view which is generally accepted nowadays. This picture leads us to the question of how a physical system like the brain can represent things. After having explained this problem, I will focus on the causal theory of representation and its two representatives in this thesis, Dennett and Dretske. I will discuss these contemporary proponents of the causal theory of representation, because they are both influential in the

philosophy of mind and language. Additionally, they both are good examples of philosophers that try to solve problems within the framework of the causal world picture, i.e. with a scientific approach.

It might seem inappropriate to combine these philosophers, because Dennett and Dretske are known to oppose and criticize each other on their views (e.g. Dennett, 1991b). However, they are part of a tradition in the philosophy of language, thought and reality that presupposes a causal view of the world, amongst which we find well known philosophers such as Russell, Ogden and Richards and Fodor (cf. e.g. Russell (1921), Ogden & Richards (1923) and Fodor (1975)). Despite their differences Dennett and Dretske do agree on a number of points, which are indisputable to them. Exactly these indisputable points are the subject of the critique that I wish develop in this thesis. My aim for this chapter is thus to distinguish the common ground that both Dennett and Dretske agree on and on which they stage their disputes. In chapter 4 this common ground will be the subject of a critical investigation with the use of Wittgenstein's philosophy, for I believe that there are a number of mistakes contained in it. Besides discussing a number of points that Dennett and Dretske agree on, I will also focus on a number of points that they both flesh out in their own way.

2.2 The Problem of Representation

Let us now try to understand why so many philosophers think the problem of representation one of the main problems in philosophy. The following story will help to illuminate this. Imagine that there is some alien life-form in the universe. Imagine that they are intelligent and that they have devised a means to visit our earth. What would they make of all the things we call representations? If they would see a picture of a person, would they understand what it is a picture of? Would they even understand that it is a picture? And what about a drawn portrait or a child's drawing? And what about written signs aside the road? Say our alien lands in a field near to a signpost that points to 'Localville'. Would he understand that it is a signpost, rather than a piece of vegetation in the field? And if he manages that, would he be able to understand what its function is or even what it says on the signpost, i.e. that it points to Localville rather than Farawayville? Initially we might be tempted to say that the signpost just points to Localville because the arrow is pointing in that direction and because it has the inscription "Localville". But on deeper reflection this becomes problematic. There is nothing in the sign 'Localville' itself that makes it a sign of *that* place rather than another place. Moreover, there is nothing in the shape of the sign that makes it point in the direction we ascribe to it, rather than the exact opposite direction. Come to think of it, we might even say that there is nothing particular about the whole signpost that enables the alien to conclude that it is supposed

CHAPTER 2 – THE CAUSAL THEORY OF REPRESENTATION

to be a representation at all, rather than e.g. a part of the natural surroundings. This does not rule out that there might be good pragmatic or biological reasons for particular types of signs to be used in a specific way. The fact that an arrow converges towards the direction it is pointing to might, for example, be used as part of a biological explanation as to why the arrow is pointing *that* direction and not the opposite one. Or the fact that the icon on the play-button of video/audio devices, which points to the right, means something like “foreward” or “go” (but not “backward”) can also be related to the fact that western cultures read from left to right and not vice versa. Furthermore, factors concerning complexity or efficiency might also prove to make certain signs more likely than others. But this can never be a sufficient explanation for *all* representations: the fact that one language uses “red” but the other “rouge” can’t be explained by such factors.

If we translate the story to our own situation, we start to wonder why *we* are able to know what representations mean. How do we know that one thing is a representation and the other not? That is, what is the difference between e.g. the word “apple” and an apple? And how do we know that “apple” refers to an apple rather than something else? This is what is so puzzling about representations: if we agree that the signpost in itself is hard to make sense of from the perspective of an alien life-form, then how can we account for our own understanding of representations? We can distinguish two problems here. One problem is how we can differentiate between representations and non-representations: how do we know that one thing is a representation, and another thing isn’t? That is, what is the difference between the sound “apple” and an apple? We want to know how it is possible in general that some things can stand in for or refer *to* other things. The second problem is that even *if* we know the difference between representations and non-representations, we still don’t know *what* a certain representation is supposed to represent. That is, given that we know that the sound “apple” is supposed to be a representation, how do we know that it is a representation of an apple, rather than something completely different? As the examples show, the symbols do not determine by themselves what they represent. There doesn’t seem to be anything particular that *makes* these representations represent the things they do rather than something else. Not for an alien, but neither for ourselves. It seems that representations do not ‘point to’ their reference intrinsically, but that there is something different going on.

In what manner, if any, are these two problems related? Will a solution for the one problem also clear the way for the other? An appealing position on this seems to be that the second problem is not really so problematic at all⁵. Once we have representations, and once they function, what specific word or sign we use to represent something involves arbitrariness to some extent: English speakers use the word “apple”, but Spanish speakers use the word “manzana” to represent the same thing: whenever a Spanish person

⁵ This is, as I take it to be, the position of causal philosophers such as Dennett and Dretske whom I will discuss in this chapter.

CHAPTER 2 – THE CAUSAL THEORY OF REPRESENTATION

wants an apple, he will ask for a “manzana”, just like an English person will ask for an “apple”. However, both languages have a sign for that specific object. Thus, on this account it seems that what particular symbol we use is thus not really the issue, as long as there is a symbol that is commonly and consistently used. The real problem seems to be how representations are possible in the first place, not how one particular representation manages to represent the thing it does. To be sure, not all representations are flexible in this manner equally. For example, if we consider pictorial representations such as photographs there obviously is a certain level of similarity needed for one thing to represent the other. However, if we proceed to the level of more abstract representations such as words and diagrams the range of possible signs will expand, until almost any sign – given that it is not too unpractical or inefficient – will suffice. So, if we are interested in the general nature of all sorts of representations as the philosophers discussed in this chapter are, an explanation in terms of similarity – either visual or otherwise – doesn’t seem to do the trick.

As a short teaser of what we will see later on, I wish to note at this point that this is one of the crucial points where Wittgenstein diverges from the philosophers discussed in this chapter. Whereas the latter try to find a general theory for all representation, Wittgenstein in fact tries to get away from the first problem by showing how particular representations manage to represent. As we shall see in Chapter 4, the relation between a word and its meaning is on Wittgenstein’s account internal, and not external. What makes e.g. ‘red’ a representation of red, is precisely the fact that we have learned to use the word in such manner, not the fact that it is a representation in the first place, which has – mysteriously – come to represent red rather than something else. In other words, whereas Dennett and Dretske try to explain the second problem by solving the first one, Wittgenstein does the opposite: he explains the representational nature of language in general by showing how particular representations manage to represent. But for now, we will have to leave the subject as it is.

Out of the two mentioned problems the first one will therefore play the largest role in this chapter: *how does one thing manage to represent another?* Or, put differently: *what makes a representation represent?* Most philosophers feel the need to solve this problem by analysing the relation between words and their referents, or between representations and what they represent (e.g., Russell, 1921; Ogden & Richards, 1923). What is it about a representation that makes it a representation? What kind of mechanism is involved? As an example of such a mechanism, *resemblance* has been considered as an explanation for representation (Crane, 1995, p. 13). Pictures and photographs, for example, seem to be rather straightforward representations of things because they resemble the things they are representing. However, resemblance turns out to be neither a sufficient nor a necessary condition for representation. It turns out to explain very little about representation (Crane, 1995, p. 17). But, even though resemblance theories have been rejected as good

explanations for representation, as an example it still highlights what type of explanation would be considered as an explanation⁶.

2.3 The Causal World Picture

Before we will consider the causal theory of representation I will in this section first discuss the causal world picture which underlies the former theory. This picture of how the world works is basically the generally accepted idea that the workings of the world are understandable in terms of causes and effects. It also holds that, in order for something to be causally effective, it has to be materialistic in the first place. In other words: non-materialistic entities such as e.g. souls cannot be said to exercise causal influence. It is the same approach that sciences such as physics and chemistry presuppose in their explanations of natural phenomena, such as the falling of objects towards the surface of the earth or the burning of wood. In order to explain such a phenomenon, we have to investigate the causal processes involved. The result of such an investigation is the specification of the relevant causes and their effects. In this way we are able to explain the phenomenon. For example: when I drop a stone, the gravitational powers of the earth will *cause* the stone to move towards the surface of the earth with increasing speed (the *effect*). The causal world picture also involves the idea that the way things work in the world is measurable and thus predictable. If we collect enough data we can describe the behaviour of these phenomena with the help of formulae and calculate, for a given case, what the outcome will be. So, for example, we are able to predict the speed, impact and falling-time of a stone if I drop it from a given height.

The causal world picture means that in a way, we think of the world as a kind of mechanism: if you set the right things in motion, the rest will follow by itself in a lawful, consistent way – and thus will be predictable. It is the idea that all natural phenomena can be measured and understood in terms of laws of nature. Crane calls this view “the mechanical world picture” (Crane, 1995, p. 2-5). As he describes it: “To put it very roughly, we can say that according to the mechanical world picture [...] things do what they do because they are caused to move in certain ways in accordance with the laws of nature” (Crane, 1995, p. 3). And it is the role of science to discover or describe these laws of nature. Even the ‘mystery of life’ itself has now been explained by science in terms of causal processes: we have discovered DNA and the chemical processes that enable organism to reproduce (cf. Crane, 1995, p. 4).

Analogously, it seems that the problem of representation should be explainable in this way as well. Since the brain is a physical structure itself, we should also be able to

⁶ See Crane (1995, pp. 13-20) for the whole range of arguments against resemblance theories of representations.

CHAPTER 2 – THE CAUSAL THEORY OF REPRESENTATION

explain the way it works in terms of cause-effect relations, i.e. in terms of natural laws. However complicated the brain itself is, it still has to work according to the laws of nature. If we combine the problem of representation and the causal world picture in this way, it leads us to the following problem: how is it possible that a physical system like the brain, which should in principle be explainable in terms of cause-effect relations, is able to represent things? Or more general: how can we give a causal explanation of representation? Any theory that does not fit within the causal world picture is by default not a candidate for the explanation of the problem of representation. For example, a theory such as Descartes dualism that postulates a non-physical mind is in direct contrast with the materialism of the causal world picture. If we want to explain the problem of representation within the framework of the causal world picture, we will have to do so in terms of materialistic causes⁷.

In the 1920's, Russell (1921) and subsequently Ogden and Richards (1923) set out to propose such a causal theory that was supposed to explain the nature of representations. For example, Russell writes that the relation between a word and its object (its meaning) "is of the nature of a causal law governing our use of words and our actions when we hear it used" (Russell, 1921, p. 198). Likewise, Ogden and Richards aim to study the use in language of signs to indicate things other than themselves. They distinguish between the sign itself (e.g. a signpost), the reference (e.g. the place the signpost is indicating) and the interpretation someone is making of the sign, i.e. his thoughts (e.g. "this sign tells me which way Amsterdam lays"). The relation between signs and their meaning is conceived in the following way: "The effects upon the organism due to any sign, which may be any stimulus from without, or any process taking place within, depend upon the past history of the organism, both generally and in a more precise fashion." (Ogden & Richards, 1923, p. 52). The main point this causal theory was that each use of a sign will eventually have some features in common, which after sufficient exposure/training we will be able to pick up. They give the following example to clarify this: you have learned to expect a flame whenever you strike a match. This is because in the past you have always experienced a flame whenever you stroke a match, i.e. whenever you made certain muscle movements, heard certain scraping sounds and had certain visual impressions. There is a causal relation between two events: the striking of a match (A) and the appearance of a flame (B). When we have accumulated sufficient experience, we can *infer* the causal relation between A and B and come to expect a flame whenever a match is stroked. In this way a thought or word is directed towards flame, i.e. it is about flame, whenever the situation is similar to other situations in which we thought about flames:

⁷ The question remains however, whether it is necessary to explain the problem of representation within the causal world picture. Does accepting the causal world picture exclude different types of explanations? What if giving an explanation of representation is a whole different *type* of explanation than explaining a (physical) phenomenon? This will be one of the points of my critique in Chapter 4.

“when a context has affected us in the past, the recurrence of merely a part of the context will cause us to react in the way in which we reacted before.” (Ogden & Richards, 1923, p. 53). And just as we learn to expect fire whenever we strike a match, we also learn to respond in the correct manner to other signs: the response we exhibit whenever we see the word “cat” is caused by our past and current experiences with this word. The same will also hold for the use of words: because every time when we see a cat we have heard (or have uttered) the word “cat” we will say “cat” in a new situation as well. Eventually, the mutual causal relation between the experience of a cat and thought or word “cat” is so strong that the whole process becomes automated and unconscious.

But, understanding language is an elaborate process of gathering evidence, inferring things from this evidence and arriving at conclusions about what the speaker/writer meant. In learning language, we supposedly become quite apt in this as described above and become able to infer conclusions ‘immediately’. As Ogden and Richards describe, “[n]ormally, whenever we hear anything said we spring spontaneously to an immediate conclusion, namely, that the speaker is referring to what we should be referring to were we speaking the words ourselves.” (Ogden & Richards, 1923, p. 15). Understanding another human being is, on their view, always a matter of arriving at the correct ‘interpretation’, something which undoubtedly goes wrong many times. That is why Ogden and Richards hold that language as a symbolic apparatus is liable to “incompleteness and defect” (Ogden & Richards, 1923, p. 19).

2.4 The Causal Theory of Representation

In this section I will discuss a number of terminological issues concerning notions such as ‘intentionality’, ‘reference’ and ‘aboutness’. Most importantly, I will discuss the views held by Dennett and Dretske about representation. I have placed both philosophers under the umbrella of the causal theory of representation, since the questions they ask and the answers they provide are similar on a number of points. Although they differ on a number of points with the abovementioned philosophers, Dennett and Dretske can still be said to be part of the same causal strand of philosophy of language. They are the heirs of the causal theories proposed by Russell (1921) and Ogden and Richards (1923). For example, the basic question Dennett and Dretske ask about representation is: “How can a mere mechanism, such as the brain, be capable of representations?”. Since the brain is just a physical system that is subject to causal laws, everything that goes on in the brain must be deterministic in the sense that everything works according to the laws of nature. Basically, if we would know the exact (causal) workings of the brain and we would be able to measure all the relevant variables (such as action potentials etc.) we could at a given time predict the future behaviour of the brain. The point is that we should somehow

be able to explain how the brain is able to represent things by giving such a causal explanation.

2.4.1 Representation and Intentionality

Representation in general means that one thing refers to, is about, or stands in for something else. If something is a representation, it has a referent, i.e. that which it refers to. The representation itself is often spoken of as being the symbol. The connection between a symbol and its referent (or meaning) is often seen as a relation. Mental states (e.g. beliefs, desires, thoughts), linguistic structures (e.g. words, sentences), but also pictures, photographs, diagrams can all bear such relations to a referent. The problem of representation is to explain this relation between the thing representing (the symbol) and the thing represented (the referent)⁸.

Representation is often discussed together with intentionality. The word ‘intentionality’ has a medieval Scholastic origin, and stems from the Latin verb *intendere*, which means being aimed, directed or extended towards something⁹. It covers the intuitive idea that mental states, i.e. beliefs, thoughts, ideas, etc., refer to something: if you think about your dog, your thought is about your dog, it refers to (or: extends towards) your dog. The term was rehabilitated by the philosopher Brentano in the nineteenth century, who held that intentionality was a mark of the mental. This is known as the Brentano thesis and holds basically that all and only mental phenomena (beliefs, thoughts, desires, etc) have intentionality. The fact that words and pictures have a meaning is, according to Brentano, only because they derive their meaning from the mental states associated with them. Thus, roughly stated, a picture represents your mother because you *think* about your mother when seeing it. Intentionality is thus easily associated with mental states such as thinking, believing, wanting or desiring (something).

Even though a number of modern philosophers hold that intentionality is not a defining feature of the mental only, I shall refrain as much as possible from using this term because it has the ‘mental’ connotation. We could say that things that have ‘intentionality’ are a subclass of things that can be called representations: ‘intentionality’ sometimes applies only to mental states, whereas ‘representation’ applies to everything that represents something else in general, such as words, signs, pictures, diagrams, and

⁸ It should be noted at this point that it is typical for the causal theory of representation to identify representation with having a referent. However, we employ other types of representations as well that are much harder to see in terms of symbol-referent relations. What about the word “hello” for example or the order “move your arm upwards”? It is by stressing the diversities of language that Wittgenstein criticizes the oversimplification of such a viewpoint, as we shall see in Chapter 3 and 4.

⁹ It should not be confused with the more common term ‘intentional’, which roughly means ‘on purpose’. Intentionality is about meaning and representation, not about volition (the will).

CHAPTER 2 – THE CAUSAL THEORY OF REPRESENTATION

the like. However, because the philosophers discussed in this chapter do not follow this distinction I will avoid it as well. When used in the subsequent text ‘intentionality’ is meant as ‘being a representation’ without the connotation of also having to be a mental state of some sorts. That is, words, pictures, diagrams and the like can also be ascribed intentionality. We are interested in representation in general, not the representation associated with mental states only. This means that we shall leave open – at least for now – whether the representational powers of certain types of representations are derived from other types of representations. As is the case, some philosophers in fact claim that Brentano had it the wrong way around: the representational powers of mental states would be derived from ‘lower-order’ representational systems¹⁰.

Before discussing the views of both Dennett and Dretske, a short explanation remains about why I have chosen to characterize their views as “the causal theory of representation”. Friedrich Waismann, who has had numerous conversations with Wittgenstein and has written substantially on Wittgenstein’s philosophy (cf. e.g. Waismann, 1979), calls a similar view attributed to Russell and Ogden and Richards “the causal interpretation of language” (Waismann, 1965, pp. 111-128). In spite of this, I chose to characterize this view rather as “the causal theory of representation” for a number of reasons. First of all, calling it a ‘theory’ conveys the characteristic of this view to try to *explain* the problem of representation scientifically, i.e. by means of proposing a hypothetical mechanism and verifying this against data (the hypothetico-deductive method). The philosophers discussed shortly both have developed a substantial theoretical model in order to account for representation. Furthermore, saying that it is about ‘representation’ rather than language conveys that this view is about the ‘new’ issue of representation in general, rather than the ‘old’ issue about how language connects up with reality, i.e. about how words and sentences correlate or reach out to the situations and objects in the world, the issue Russell and Ogden and Richards were concentrating on.

I will now discuss the views of both Dennett and Dretske on a number of matters that concern, in their view, the nature of representations. First I will discuss the viewpoint of Dennett by answering a number of questions that give a sufficient idea of his theory about representations. Subsequently, I will discuss the answers to the same questions that Dretske has provided in his writings.

¹⁰ This is in fact what Dennett claims. He writes, talking about self-replicating macromolecules: “These impersonal, unreflective, robotic, mindless little scraps of molecular machinery are the ultimate basis of all the agency, and hence meaning, and hence consciousness, in the world.” (Dennett, 1996, p. 29). Dretske disagrees: he holds that at least some representations derive their representational powers from *us*, such as when we use coins to represent the players of a basketball game. He calls such systems Representation Systems of Type I or II (cf. Dretske, 1988, Chapter 3).

2.4.2 Dennett

2.4.2.1 *What is representation?*

Dennett¹¹ equals intentionality with ‘aboutness’. According to Dennett, aboutness is the defining feature of a representation: everything that is a representation is about something. He writes: “Something exhibits intentionality if its competence is in some way *about* something else. An alternative would be to say that something that exhibits intentionality contains a *representation* of something else – but I find that less revealing and more problematic.” (Dennett, 1996, pp. 46-7). He puts emphasis on the intuitive distinction between things that have aboutness and things that lack aboutness: a picture of your mother is about your mother, a book about Amsterdam is about Amsterdam, and the word “apple” is about an apple. But your mother, Amsterdam and an apple are not about something else.

The problem, according to Dennett, is to explain how aboutness is possible in terms of causes and their effects. His general aim in philosophy is “to figure out how the brain could possibly accomplish the mind’s work ... how the mechanical responses of ‘stupid’ neurons could be knit into a fabric of activity that actually discriminated meanings” (Dennett, 1994, p. 236). Dennett is most intrigued by the fact that something that is so mechanical and physical as the brain is able to exhibit aboutness, i.e. is able to represent the things around itself. Dennett is thus looking for a theoretical explanation of how the brain is able to exhibit representations *within* the framework of the causal world picture that we discussed in the previous section: we are looking for an account that will provide us with an explanation of representation in terms of causes and effects. He writes that he sees it as his task to sketch “the outlines of a physical structure that could be seen to accomplish the puzzling legerdemain of the mind” (Dennett, 1994, p. 237). In order to be able to explain consciousness – which Dennett conceives as one of the two main topics of the philosophy of mind (the other being intentionality) – we first have to be able to explain intentionality, which is “more fundamental than consciousness” (Dennett, 1994, p. 236). That Dennett holds the causal world picture to be the absolute basis for his explanations is obvious: “When mechanical push came to shove, a brain was always going to do what it was caused to do by its current, local, mechanical circumstances, whatever it *ought* to do.” (Dennett, 1994, p. 237). Dennett even explicitly decrees in *The*

¹¹ It has to be noted at this point that Dennett himself claims to be a follower of Wittgenstein (Dennett, 1991a, p. 463). Others however, do not agree: “Dennett’s accounts of intentionality and consciousness are very far indeed from anything Wittgenstein would have countenanced.” (Bennett & Hacker, 2003, p. 413) I agree with Hacker on this point and will thus not see Dennett as a follower of Wittgenstein.

Intentional Stance: “I declare my starting point to be the objective, materialistic, third-person world of the physical sciences.” (Dennett, 1987, p. 5).¹²

2.4.2.2 What kind of things can be called a representation?

To Dennett, aboutness is not restricted to the realm of the mental at all, as it was for Brentano. He writes about thermostats, amoebas, opinioid receptors in the brain, locks and keys, chess computers – amongst others – that they all exhibit aboutness (Dennett, 1996, p. 45; p. 47; Dennett, 1987, p. 31). The common characteristic of all the entities he ascribes aboutness to, is that they can be called an *intentional system*. Dennett defines intentional systems in turn by referring to the fact that they are all predictable from the intentional stance (Dennett, 1996, p. 45). We shall therefore have to explore his idea of the intentional stance first, in order to comprehend Dennett’s views on representation.

Dennett devised the idea of the *intentional stance* as a description of the strategy of interpretation that we as humans use in order to explain or predict the behaviour of certain systems around us. He contrasts it to the *physical* and the *design* stance. The physical stance is the way we explain the phenomena of nature as inanimate objects, such as the way physics describes a falling stone. The design stance is the way we expect things to work as they are designed to do, e.g. we expect an alarm clock to go off at 7:00 if we have set it to do so. The intentional stance is a way of interpreting entities *as if* they are rational, intentional agents that have beliefs, desires, thoughts – i.e. mental states. That is, according to Dennett we adopt a certain stance in order to explain the behaviour of other human beings, but also that of less ‘conscious’ systems such as a chess playing computer or a thermostat. Whenever we explain such entities behaviour by using descriptions such as ‘it thinks that it is too warm in here’ or ‘it believes that such-and-such a move in the game will increase its winning chances’ we are adopting this stance and (successfully) predicting or explaining the entities behaviour.

However, Dennett takes this viewpoint not just as a convenient way of talking about things. If we assume that X is an intentional system it might seem that we are not sure about the fact whether it is *really* intentional or not, for how have we learned what intentionality really is? However, Dennett believes, if we can explain or predict the behaviour of a certain system with the use of the intentional stance, it thereby *is* an intentional system: “*Intentional systems* are, by definition, all and only those entities

¹² Dennett seems to be under the impression that philosophical problems are solvable “through a combination of scientific enquiry and the adjustment of our conceptual prejudices in light of empirical evidence” (Symons, 2002, p.12; the book is not written by Dennett but it appears to have his imprimatur). Again we see that Dennett adheres to the causal world picture (and the natural sciences). Dennett is not always as clear on this point as here, but seen that these are recent publications, I will take it that he does in fact work against the background of the causal world picture.

CHAPTER 2 – THE CAUSAL THEORY OF REPRESENTATION

whose behaviour is predictable/explicable from the intentional stance” (Dennett, 1996, p. 45)¹³. Moreover, Dennett also claims that “when we discover some object for which the intentional strategy works, we endeavour to interpret some of its internal states or processes as internal representations” (Dennett, 1987, p. 32; original in italics). In a sense, says Dennett, the lock is about the key and the thermostat is about the room temperature: in both cases the one thing can tell us something about the other. So even if we consider a lock and key, a thermostat or a chess-playing computer, we can use the intentional stance with success. And therefore we can call these entities an intentional system and call their internal states representations.

Dennett’s idea of the intentional stance is strongly based on his idea that folk psychology is a naïve form of science. When in everyday situations we take on the *intentional stance* towards other human beings, we *assume* that others are intentional systems: we explain and predict their behaviour as the result of intentional states such as reasons, beliefs, desires or fears that they have. We can then see their internal states *as* references, i.e. we can see their internal states as if they are about other things (as if they have aboutness), because we have a psychological model of other (human) beings. According to Dennett, ‘folk psychology’ is a theory or theoretical model about other (human) beings that enables us to interpret and predict their behaviour. We as humans formulate hypotheses about what other people mean or about what they will do based on the information we have, i.e. external, behavioural information. This model enables us to make reliable and accurate predictions and explanations of other peoples behaviour: “We use folk psychology – interpretation of each other as believers, wanters, intenders, and the like – to predict what people will do next”, and our power to interpret the actions of others is dependent on the power to predict them (Dennett, 1991b, p. 29). Therefore, Dennett believes that folk-psychology is a useful source of theory (cf. Dennett, 1987, pp. 43-57). We are in effect always *interpreting* the behaviour (and utterances) of other humans: “folk-psychology might best be viewed as a rationalistic calculus of interpretation and prediction – an idealizing, abstract, instrumentalistic interpretation method that has evolved because it works and works because we have evolved.” (Dennett, 1987, pp. 48-49). The fact that we understand other humans is, on this view, a result of having arrived at correct interpretations¹⁴, based on a very intricate model that has evolved during the course of history: folk psychology. And the success of this

¹³ Cf. Dennett (1987), p. 15. Dennett here discusses the intentional stance with respect to being what he calls a true believer: “any object – or as I shall say, any *system* – whose behavior is well predicted by this strategy is in the fullest sense of the word a believer. *What it is* to be a true believer is to be an *intentional system*, a system whose behavior is reliably and voluminously predictable via the intentional strategy.”

¹⁴ Remind that what Ogden and Richards have claimed presented understanding as a form of interpretation as well, see 2.3 (final paragraph).

CHAPTER 2 – THE CAUSAL THEORY OF REPRESENTATION

primitive scientific model is dependent on “there being some order or pattern in the world to exploit” (Dennett, 1991b, p. 30) – a pattern which we can discern in the behaviour of other intentional beings.

To sum up, because it has had evolutionary advantages, we have developed talk about beliefs, desires, thoughts and the like because this gives us the power to predict and understand the behaviour of other beings. This talk is not merely a convenient way of describing, it has ontological status as well: if the intentional stance can be successfully applied to a system, we are in virtue of that characteristic justified in calling it a true intentional system.

2.4.2.3 *How does representation work?*

Dennett talks about the ‘internal state’ of simple intentional systems as being *representations* of their surroundings. For example, if we consider a thermostat that measures the room temperature and sends signals according to its measurements to a boiler, we can say that the thermostats internal states are in fact representations:

There is a familiar way of alluding to this tight relationship that can exist between the organisation of a system and its environment: you say that the organism continuously mirrors the environment, or that there is a representation of the environment in – or implicit in – the organisation of the system. (Dennett, 1987, p. 31).

Thus, whenever a system has an internal state that can be said to mirror, or represent, the environment, we can call this a representation. Moreover, Dennett speaks about such representations – certain internal states of a given system – as having causal control over the ‘actions’ of the system. The thermostat, for example, will notice changes in the environment and will accordingly influence the boiler with its signals (cf. Dennett, 1987, p. 31). He writes: “What makes some internal feature of a thing a representation could only be its role in regulating the behaviour of an intentional system” (Dennett, 1987, p. 32; original in italics). Later, he writes accordingly: “*beliefs* are information-bearing states of people that arise from perceptions and that, together with appropriately related *desires*, lead to intelligent *action*.” (Dennett, 1987, p. 45). Clearly, he thinks that mental states (and representations) have causal power, i.e. they bring about behaviour¹⁵. Dennett stresses that there is no reason to think of ourselves as different from the thermostat. Even though we are more complex and have more complex representations, the basic

¹⁵ The quote is Dennett’s description of folk-psychology, but as noted before, Dennett believes this to be a very reliable model for human behaviour (as part of the intentional stance). Moreover, directly after the quote he concludes: “That much is relatively uncontroversial”.

CHAPTER 2 – THE CAUSAL THEORY OF REPRESENTATION

mechanism of representation is similar. There is only a difference in degree of intentionality between us and the thermostat.

In fact, Dennett encourages us to think of ourselves as being composed of ‘robots’ or ‘automata’, such as the macromolecules that make up our DNA, hemoglobin molecules, antibodies, neurons etc. (Dennett, 1996, p. 30). All these simple automata have no more than a very basic functionality, yet, together they make up our bodies. Our nervous system is composed of neurons, which can in turn be seen as composed of a number of even smaller automata. Thus, Dennett concludes: “An autonomous nervous system is not a mind at all but rather a control system, more along the lines of the nutritive soul of a plant, that preserves the basis integrity of the living system.” However, if we compare simple automata to our own mind, we should come to see that there is no significant difference:

We sharply distinguish these ancient systems from our minds, and yet, curiously, the closer we look at the details of their operation the more mindlike we find them to be! ... It is as if these cells and cell assemblies were tiny, simple-minded agents, specialized servants rationally furthering their particular obsessive causes by acting in the ways their perception of circumstances dictated. (Dennett, 1996, pp. 34-35).

Thus, instead of the vague notion of the mind we can, on Dennett’s account, also conceive of the embodied human brain that it is a rather complex control system that has evolved over millions of years to end up functioning the way it does now: using representations of the outside world to govern its behaviour in order to survive.

2.4.3 Dretske

2.4.3.1 *What is representation?*

Dretske holds a similar view to Dennett. He writes: “Intentional states (and, therefore, cognitive states) appear to have something like *meanings* (propositions) as their object (content), as that *on which* the mind is directed” (Dretske, 1980, pp. 354-5). Things that exhibit aboutness – intentional states in Dretske’s book – are about something else and can be said to have a ‘content’ (i.e. that which they are about). Dretske sees a mental state as an internal physical state of the brain that has managed to gain control over behaviour, which in turn is seen as the output of the system: “Experiences and beliefs are merely those internal, presumably physical, states of a system having the function of providing information (in the case of experience) and mobilizing it (in the case of belief) for use in the control of behaviour.” (Dretske, 1994, p. 260). Thus, like Dennett Dretske believes mental states (and representations) to exhibit causal powers: they bring about our behaviour.

CHAPTER 2 – THE CAUSAL THEORY OF REPRESENTATION

The problem of representation according to Dretske, is like Dennett's, to try and make sense of representation within the framework of the causal world picture. Dretske writes: "Some account *must* be given of how a purely physical system could occupy states having a content of this sort" (Dretske, 1980, p. 355; my italics). Like Dennett, Dretske is looking for a theoretical explanation that is able to account for meaning in the brain, taking the latter as a physical structure that can occupy states that exhibit aboutness or representation. He also conceives of mental states as having causal powers: "My reasons, my beliefs, desires, purposes, and intentions, are – indeed they must be – the cause of my body's movements." (Dretske, 1988, p. ix). The aim of his philosophical doctrine about representation is "to see how reasons – our beliefs, desires, purposes, and plans – operate in a world of causes, and to exhibit the role of reasons in the *causal* explanation of human behaviour" (Dretske, 1988, p. x). This shows that we can safely assume that Dretske, like Dennett, is operating from the background of the causal world picture.

2.4.3.2 *What kind of things can be called a representation?*

Like Dennett, Dretske thinks that intentionality is not restricted to the mental. If we approach the problem from the viewpoint of simple communication systems, such as thermostats or galvanometers, "it soon becomes clear that intentionality, rather than "being a mark of the mental," is a pervasive feature of *all* reality – mental and physical. Even the humble thermometer occupies intentional states." (Dretske, 1980, p. 356). The difference between the cognitive states of humans, such as beliefs, and the intentional states of simple communication systems, is simply the fact that the former have a "higher order of intentionality" (Dretske, 1980, p. 356).

Like Dennett, Dretske believes a thermometer to be *about* the temperature of the medium it is in: it tells us something about the physical state of that medium. Nonetheless, a thermometer doesn't *know* what the medium is – it only shows us its temperature (Dretske, 1980, pp. 358-9). Whether we put it into water, steam or oil, it will simply tell us what the temperature of each of these substances is. To account for this difference between a thermometer and cognitive systems, Dretske distinguishes between representations in general and cognitive states in particular. Cognitive states are also representations, but they are of a "higher level of intentionality" (Dretske, 1980, p. 358). Unlike Dennett, Dretske thinks it is incorrect to predict or explain the behaviour of a thermometer or chess-playing computer using descriptions such as "believing it is so-and-so many degrees Celsius" or "thinking that he is winning". We are not justified to ascribe knowledge to a thermometer. About a similar instrument Dretske writes: "Despite the fact that galvanometers receive, process, and display ... information about affairs external to them, they do not occupy cognitive states." (Dretske, 1980, p. 358).

CHAPTER 2 – THE CAUSAL THEORY OF REPRESENTATION

Such systems do however, according to Dretske, represent the affairs external to them. This is because cognitive states require a higher level of intentionality than simple systems such as thermometers have: a thermometer cannot distinguish between measuring the temperature of water and measuring that of oil. As said, it will simply ‘represent’ the temperature of whatever medium it is immersed in. If we consider a different system that has measuring systems both for measuring the temperature and e.g. for measuring the structure of the medium it is immersed in, we can see the beginnings of a system that is able to know something: that it is *water* that has a certain temperature. Such a system can know at the same time that something is 20 degrees Celsius and that something is water, while an ordinary thermometer can only ‘know’ that something is 20 degrees Celsius without being able to separate this from the fact that it is water that has that temperature. In other words: it cannot carry the information that there X is 20 degrees Celsius without carrying the information that there is water¹⁶. Therefore, it cannot be said to ‘know’ anything at all. In contrast to this, the new and improved system *will* be able to carry the information that X is water without carrying the information that X is 20 degrees Celsius precisely because it has a separate and independent ‘detector’ for water. Thus, cognitive states like knowing only come in when a system is complex enough to code information in different ways about one and the same entity. In that manner, a system’s internal states will reflect the cognitive differences between “this is water” and “this is 20 degrees Celsius”. The more detectors a system has, the higher its degree of cognition is likely to be. Which detectors are important is a matter of evolutionary selection: all external information that has been important for the survival of e.g. our species is in some way represented by the different receptors we have to detect it.

2.4.3.3 *How does representation work?*

The way Dretske explains for representation is as follows: he conceives of meaning as a nomic relation between a source of information, the referent, and a receiver of information, the representation. Whenever there is a nomic, i.e. lawful, relation between a certain (internal) state of a system on the one hand and a certain state of the environment, there is communication. If we consider the example of the thermometer, we can explain this in more detail: a thermometer indicates the room temperature precisely because there is such a nomic relation between the (internal) state of the thermometer and the room temperature. Every time when something changes in the room temperature, the state of the thermometer will change accordingly. And because this is a *lawful* regularity, we can also predict what will happen to the state of the thermometer if we change the room

¹⁶ Note that on Dretske’s account an ordinary thermometer cannot *know* anything because it carries too much information: a thermometer is thus said to carry the information that “the water it is immersed in is so-and-so many degrees Celsius” and not “X is so-and-so many degrees Celsius”.

CHAPTER 2 – THE CAUSAL THEORY OF REPRESENTATION

temperature accurately. Here again we see a similarity with Dennett, for the latter believes that our power to *interpret* the actions of other people depends on our power to *predict* them, (Dennett, 1991b, p. 29). Dretske consequently writes:

Any physical system, then, whose internal states are lawfully dependent, in some statistically significant way, on the value of an external magnitude (in the way a properly connected measuring instrument is sensitive to the value of the quantity it is designed to measure) qualifies as an intentional system. (Dretske, 1980, p. 357).

Dretske also provides for a way to distinguish between knowledge and representation. A ‘simple’ representational system, such as a thermometer, doesn’t *know* anything because it cannot distinguish between measuring the one medium from measuring the other. It will simply convey the information of the temperature to the reader of the thermometer, indifferent to the medium it is immersed in. When a certain state indicates some environmental feature, say T, it need not necessarily also contain other information about that feature, e.g. that it is also S or that it is W that has T. For example, a thermometer can represent that something is 24 degrees, without representing that it is in fact water (or oil) that is having that temperature. Cognitive states, like our beliefs and desires, have a higher degree of intentionality because they are able to discriminate such information. Since we have evolved to represent a very large quantity of information about our environment, most of our representations do involve such information, because we have several ‘measuring instruments’ working at the same time. For example: when we immerse our elbow in water we not only measure the temperature but we also feel the structure of the fluid which tells us that it is water and not oil.

Dretske furthermore holds that “the intentionality of our cognitive states has its source in the intentionality of informational structures”. Their intentionality is derived from the fact that cognitive states are nomically dependent on other states, either environmental or internal (Dretske, 1980, pp. 357-8). He furthermore writes: “what makes some brain states into mental states (experiences or beliefs, as the case may be) is related to their information-carrying function. Just as information can convert a belief into knowledge and an experience into perception, the function of providing (in the case of perception) and using (in the case of belief) such information can convert physical states into mental states: an experience or a belief.” (Dretske, 1994, p. 260).

Dretske is more subtle in his account of representations than Dennett. He distinguishes between different sorts of representations in his book *Explaining Behaviour*. The first type of representations, RS type I, derive all their representational powers from us, for example when we use pins to represent the positions of armies in the strategy game Risk: these pins do not represent armies intrinsically, but only because we make them represent the armies in our game (Dretske, 1988, pp. 52-4). Note that in this

CHAPTER 2 – THE CAUSAL THEORY OF REPRESENTATION

case there is indeed no lawful regularity between actual armies and the pins. The second type of representations, RS type II, *do* have a power to indicate that is independent of us, but which representation they actually have is still dependent on the function we assign to them (Dretske, 1988, pp. 54-62). For example: a thermostat could also be fixed to the controls of a train in such a manner that it keeps the train at a certain speed, instead of keeping the room temperature at a certain level. Nonetheless, a thermostat is still said to have a representation. Another example is an electrical fuel gauge of a car, which indicates both the level of fuel left and the amount of electrical current that flows through the wires connecting the gauge to the tank – the latter is in fact the means to do the former, but since it is its function to measure the fuel and not the electricity, *that* is what we say it represents (Dretske, 1988, p. 59). Fixed in a different situation it could, by means of its functionality as an electrical-current measurer, be made to represent something different.

However, in contrast to RS type I and II the third type of representation, RS type III, is not dependent on anything other than itself: its powers to represent are *intrinsic*. They indicate something no matter what, quite unlike the pins in the game of Risk. These representations are found in what Dretske calls natural systems of representation (Dretske, 1988, pp. 62-64). Dretske considers as an example the senses of an animal: “An animal’s senses (at least the so-called exteroceptors) are merely the diverse ways nature has devised for making what happens inside an animal depend, in some indicator-relevant way, on what happens outside.” (Dretske, 1988, p. 62). Basically, just like a thermometer’s state is dependent on the room temperature, certain internal states of animals will be dependent on – and thus indicative for – external situations. That dependency between the two states is what representation consists of. Dretske writes:

“Just as we conventionally give artifacts and instruments information-providing functions, thereby making their activities and states – pointer readings, flashing lights, and so on – representations of the conditions in the world in which we are interested, so learning converts neural states that carry information – ‘pointer readings’ in the head, so to speak – into structures that have the function of providing some vital piece of the information they carry.” (Dretske, 1994, p. 261).

Two things are noteworthy in what Dretske writes here. The first thing is that he calls the activities and states of instruments, such as thermometers, ‘representations’ of conditions in the world. The second thing is that he calls neural states ‘information providing structures’, meaning that neural states *contain* or even *are* representations of something other than themselves.

CHAPTER 2 – THE CAUSAL THEORY OF REPRESENTATION

“The later Wittgenstein, in my view, has no ancestors in the history of thought. His work signals a radical departure from previously existing paths of philosophy.”

- Georg Henrik Von Wright

Chapter 3

Wittgenstein's Method

3.1 Introduction

Although Wittgenstein published only one work in his lifetime, there is a good amount of his writing published at the moment, ranging from prepared publications to notes taken by students attending his lectures. Moreover, the corpus of secondary literature is impressive, to say the least, showing that there is great interest in what Wittgenstein had to say. Moreover, the issue of representation presents one of the most important topics in his work. As noted before, Wittgenstein sets out to show errors in philosophy resulting from a poor understanding of the forms of our language. But, as he notes himself, we can only understand such errors truly if we have made them ourselves. Most philosophic errors that Wittgenstein tries to amend are thus errors he himself has made as well, or to which he has succumbed to at one time or another:

One must start out with error and convert it into truth. That is, one must reveal the source of error, otherwise hearing the truth won't do any good. The truth cannot force its way in when something else is occupying its place. To convince someone of the truth, it is not enough to state it, but rather one must find the path from error to truth (RF, p. 61).

The problems we are thus facing in this thesis are problems that Wittgenstein himself also found hard to think about the right way. Wittgenstein's insights into the problem of representation are thus the result of his own entanglement with this problem. In order to understand Wittgenstein's ideas on representation and meaning, in contrast with the causal theory of representation, it is fundamental to make sense of both his method of doing philosophy and his view on the role of this ancient discipline. In this chapter I shall therefore provide a short overview of Wittgenstein's philosophical style; of his view on the role of philosophy; of his method of solving philosophical problems – called a

grammatical investigation – and its characteristics; and finally of his main instruments for conducting grammatical investigations, such as the concepts of family resemblance and language game¹⁷.

3.2 Wittgenstein's style

Wittgenstein's style of writing is highly distinctive within the whole of philosophical literature. Instead of dividing his work in chapters, sections and paragraphs, his writings consist generally of ordered remarks varying in length from a single line to several paragraphs¹⁸. But the form of Wittgenstein's writing is not the only thing that stands out against other great philosophical works. His remarks often do not seem to present an explicit line of argument, let alone clearly stated conclusions. Many remarks revisit subjects again and again without ever seeming to provide an explicit and final statement on the subject at issue. Moreover, sometimes they are presented as a discussion between Wittgenstein and an interlocutor, and it is not always clear which remark is to be ascribed to whom. Questions that are posed in the remarks are often left unanswered, analogies are sometimes very hard to understand and thought-experiments seem to be highly unlikely or simply impossible.

Overall, these distinctive features of Wittgenstein's writing make it exceedingly difficult to understand his work. But, to Wittgenstein himself they are intrinsic to his method of doing philosophy. The point of many remarks is obscure on the first reading, and it is often unclear whether Wittgenstein agrees or disagrees with a view that is under investigation. What his own views are can only be revealed after careful interpretation of his remarks. And in order to interpret these remarks, we have to understand Wittgenstein's method of and view on philosophy.

That Wittgenstein's method of philosophy and his style of writing are intricately related is something he himself discusses in the preface to the *Philosophical Investigations*. He describes his remarks as the result of his own philosophical investigations; as the recordings of his thoughts. That this did not result in a fluid and transparent line of argumentation is something Wittgenstein clearly acknowledges himself:

¹⁷ For a general overview of Wittgenstein's style and method see McGinn (1997) and Savickey (1999).

¹⁸ *The Blue and Brown Books* form an interesting exception here. In this work Wittgenstein has elaborated his remarks into a more standard form of fully written paragraphs and he is generally much more explicit in stating his point.

The best that I could write would never be more than philosophical remarks; my thoughts were soon crippled if I tried to force them on in any single direction against their natural inclination. – And this was, of course, connected with the very nature of their investigation. For this compels us to travel over a wide field of thoughts criss-cross in every direction. (PI, preface)

As we will see – and can already catch a glimpse of in the above quote – it is the anti-theoretical and anti-generalising nature of Wittgenstein's investigations that demands for this approach. Wittgenstein also compares the *Philosophical Investigations* to an album delivering different sketches from different viewpoints of the same landscape. This way of approaching the philosophical problems Wittgenstein wishes to eradicate is something that we can clearly recognise in both the method and style of his writings.

3.3 The negative role of philosophy

For Wittgenstein, language is both the source of and the solution for philosophical problems. What philosophy is to achieve, is not the production of scientific knowledge, but rather to contribute to human understanding. Philosophy is not an empirical science that deals with (the creation or discovery of) facts, but it deals with conceptual investigations. There is nothing new for philosophy to discover, its role is simply to put together what we already know concerning a subject in such a way that we can see how everything around it is connected. The reason why we fail to see these connections in the first place, is because we do not understand the form of our own language. A true philosophical investigation should concern itself with our language, clarifying the conceptual confusions that cause us to look for the wrong answers for our philosophical questions.

This section will explain the negative purpose of philosophy. It has to be noted at this point that Wittgenstein's conception of philosophy is not shared by most philosophers. His conception of philosophy turns the bulk of both traditional and contemporary philosophy into social sciences like anthropology, sociology, psychology etc. or into nonsensical considerations. Indeed, there is a vast tradition of theoretical philosophy that follows its own methods and creates philosophical theories and models, but these are mostly considered non-philosophical in Wittgenstein's view. Let us now take a look at Wittgenstein's reasons for criticizing such investigations.

3.3.1 The source of philosophical problems: the theoretical attitude

In PI 109 Wittgenstein makes a number of statements about what philosophy is – and is not: 1) philosophy is not a science; 2) there is no place for theories in philosophy; 3) philosophy must not contain hypotheses; 4) it explains nothing, only describes what is already there; 5) the problems are not empirical, i.e. new empirical data will not help in solving philosophical problems; 6) problems can only be solved by rearranging what we have always known, i.e. by getting a clear view of the way language functions (a perspicuous representation) (cf. RF, p. 69; Baker & Hacker, 1980, p. 223). Of these, points 1, 2, 3 and 5 are related to what I have called the negative role of philosophy. These points will be discussed subsequently. Points 4 and 6 will be discussed in the next section on the positive role of philosophy (section 3.4).

So why is it that Wittgenstein says that we can't find the answers to our philosophical questions by using scientific methods, such as discovering new data or constructing a theory? Wittgenstein certainly understands that we feel as if this sort of thing would be the solution for our problems:

We feel as if we had to penetrate phenomena: our investigation, however, is directed not towards phenomena, but, as one might say, towards the 'possibilities' of phenomena. (PI 90)

I feel as though, if only I could fix my gaze absolutely sharply on this fact, get it in focus, I must grasp the essence of the matter. (PI 113)

And the fact that we are inclined to think this way is exactly the problem, as McGinn says: "The real fault, Wittgenstein believes, is not in our explanations, but in the very idea that the puzzlement we feel can be removed by means of a discovery" (McGinn, 1997, p. 19). The idea that the construction of theoretical models will bring the solution is itself the obstacle. McGinn calls this the *theoretical attitude* (McGinn, 1997, p. 16). It concerns the tendency to apply scientific methods in philosophy. When we ask philosophical questions, we "try to grasp the *essence*" (PI 116) of the subject at issue. We are tempted to answer our questions in the way science does: by trying to uncover or explain the nature of the phenomena at issue. McGinn describes this as follows:

[I]n the very act of framing these questions, we are tempted to adopt an attitude towards these phenomena which, Wittgenstein believes, makes us approach them in the wrong way, in a way which assumes that we have to uncover or explain something. When we ask ourselves these questions, we take us a stance

towards these phenomena in which they seem suddenly bewilderingly mysterious [...]. (McGinn, 1997, p. 18)

This theoretical attitude towards the phenomena is not beneficial in philosophy, because the nature of philosophical problems is a misunderstanding of the question. It not a problem that lies in the nature of the phenomenon, thus, trying to find a scientific answer – i.e. some kind of theory – will not do us any good. We can almost see Wittgenstein sighing when he writes: “The learned who would always like to have *a* theory!!!” (RF, p. 81). He elaborates:

Philosophers constantly see the method of science before their eyes, and are irresistibly tempted to ask and to answer questions in the way science does. This tendency is the real source of metaphysics and leads philosophers into complete darkness. (BB, p. 18)

Instead, Wittgenstein asks us to go ‘against the grain’ of what we generally tend to do when posed with a philosophical problem: we should try *not* to approach our problems scientifically. Wittgenstein is here not stating his opinion on a philosophical subject, but is only trying to persuade us to use his conceptual method of doing philosophy (cf. WLFM 103). Wittgenstein is not convincing us to think in the opposite direction, “but his new style of thought is one that eschews the abstract theorizing which he believes lies at the root of philosophical confusion” (McGinn, 1997, p. 29).

Thus, not only does Wittgenstein see philosophy as fundamentally different from science, he also sees the application of scientific methods as one of the roots of philosophical problems. It is not that Wittgenstein opposes science or its results, but rather that he sees the way of asking questions and looking for answers in science as inappropriate for the species of problems in philosophy. The kind of ‘explanation’ that Wittgenstein offers, is to be seen as fundamentally different from the scientific explanation that he wants us to get away from. In science phenomena are explained by means of causal hypotheses, idealisations and hypothetical-deductive inferences. Such explanations have no place in philosophy, for here there are only explanations of meaning based on our everyday use of language. There is nothing to *uncover* in philosophy, as in case of an empirical question, because everything already lies open to view (PI 126). It is the role of philosophy to remind us of the ordinary meaning of words, which is presupposed in our philosophical considerations.

This conception of philosophy can be characterized as a negative approach: instead of solving the alleged problem, Wittgenstein’s investigation only amounts to the rejection of existing philosophical theses. He forbids the use of hypotheses, (scientific) explanations and theories in philosophy. So he wonders: “Where does our investigation get its importance from, since it seems to destroy everything interesting, that is, all that is

great and important?” (PI 118)¹⁹. But whoever asks this question, has not appreciated the fundamental nature of Wittgenstein’s method. He is not trying to convince us of a certain doctrine or opinion, but he wants to do away with all scientific theorizing within philosophy. Thus, he replies: “What we are destroying is nothing but houses of cards and we are clearing up the ground of language on which they stand.” (PI 118). It is the task of philosophy to uncover pieces of nonsense that have originated from philosophical misuse of our language (PI 119).

3.3.2 Misled by the form of our language: the character of depth

Nonetheless, Wittgenstein realizes that the problems of philosophy are easily embraced. They are not mere mistakes but rather superstitions or misunderstandings (cf. PI 110) about the nature of things. It is language itself that lures us into seeing things according to a wrong picture (PI 115). He writes:

The problems arising through a misinterpretation of our forms of language have the character of depth. They are deep disquietudes; their roots are as deep in us as the forms of our language and their significance is as great as the importance of our language. (PI 111)

Philosophy, as we use the word, is a fight against the fascination which forms of expression exert on us. (BB, p. 27)

Philosophy is a battle against the bewitchment of our intelligence by means of language. (PI 109)

Philosophical problems are thus not so much concerned with the nature or state of the reality around us, but are about our understanding and use of words. That is, most philosophers try to discover the *essence* of the subject they discover²⁰. They will for example compare all cases in which we speak of ‘knowledge’ or ‘being’ and then try to extract from all these examples the essence of that subject (e.g. by looking for a common feature in all the cases). Philosophical problems often seem to have an ontological or

¹⁹ Cf. Hacker, 2003, p. 332: “Indeed, for a while, in 1931, he conceived of himself as the destroyer, the *terminus ad quem*, of the great tradition of western philosophy.”

²⁰ This point has to be taken with some considerations: I feel that Wittgenstein *is* in fact discussing reality, but he does so by talking about concepts and how they are (internally) related to other concepts and our daily activities. The viewpoint Wittgenstein is giving is, in my opinion, the only rewarding way to talk about ‘reality’ as such. Hopefully the reader will come to appreciate this after reading Chapter 4.

metaphysical nature, but Wittgenstein wants to show that this is only a chimera (cf. PI 116, 117). To solve a philosophical problem, he says, one has to undertake a different sort of investigation, since constructing a new theory or discovering new facts about reality isn't going to help us out. Wittgenstein describes this type of investigation as a *grammatical* one, saying that it "sheds light on our problem by clearing misunderstandings away" (PI 90). Again and again, Wittgenstein reminds us that the real problem is in our understanding of our language, not in the structure of reality itself:

Here it is difficult to see that what is at issue is the fixing of concepts. A concept forces itself on one. (This is what you must not forget.) (PI, p. 204)

We are being misled by the form of our language, by looking at it from a wrong perspective. For example, if we compare the sentence "this is red" with "an apple is a fruit" the word 'is' cannot be explained in both cases by reference to some common feature – as would be a sufficient explanation for a physical phenomenon such as different falling objects (i.e. gravity). Another interesting example is the following: even though visible is related to vision as sensible is related to sensing (touch), desirable is not related to desire in the same way. As Wittgenstein says: "A simile that has been absorbed into the forms of our language produces a false appearance, and this disquiets us" (PI 112). In PI 89 Wittgenstein gives an example of how a simile in the forms of our language can mislead us: for when we ask "What is time" we feel as if the answer should be something similar to the answer of "What is the specific gravity of hydrogen", since both sentences are akin in structure. But this first question cannot be answered in the same manner as the second. If we assume that it *is*, we are being held captive by a picture of how things are, which language seems to force on us (PI 115). We need to start looking at such questions from a different point of view. In the Big Typescript we find the following section heading:

Difficulty of philosophy, not the intellectual difficulty of the sciences, but the difficulty of a change of attitude. Resistances of the will to be overcome. (BT 406)

It is the picture that language forces on us (PI 115) that we need to get rid of – and it is by making a change of attitude that we can do so: we will have let go of our theoretical attitude.

3.4 The positive role of philosophy

We have seen that Wittgenstein's philosophical method can be characterized in a negative way: he wants to do away with all scientific tendencies within philosophy, including the use of theories, hypotheses and explanations. However, there is also a positive purpose to his philosophy: the idea of a perspicuous representation as a means to battle the conceptual bewilderment caused by our language. This section will elaborate on this positive purpose of Wittgenstein's method.

3.4.1 Ploughing through the whole of language

Instead of succumbing to the theoretical attitude, we must try and attack the root of our problems. One must "plough through the whole of language" (RF, p. 68) because the real problems lie already in the very first steps of thinking about the phenomena that puzzle us. We should conduct a *grammatical investigation*: examine our use of the concept at hand and the conceptual relations that are involved. Instead of looking for new knowledge, we should look for a better understanding of the conceptual relations that lie before us in our own language. Only then will we be able to understand what it does and does not make sense to say.

Wittgenstein's method consists in clarifying how the involved concepts actually function in everyday life. We need to compare the concepts at issue with related concepts, put them at work in concrete examples and remind ourselves of how we actually use them. For it is in language at work that we see the relations between concepts:

The confusions which occupy us arise when language is like an engine idling,
not when it is doing work. (PI 132)

Wittgenstein reminds us that "one must always ask oneself: is the word ever actually used in this way in the language-game which is its original home?" (PI 116). For most philosophical problems arise when we think about words *outside* of their actual use, when we try to give them a special, metaphysical, meaning inside philosophy. But we must understand that the sense of a word is not something that simply sticks to it, no matter what we do with the word. Wittgenstein retorts: "[w]hat we do is to bring words back from their metaphysical to their everyday use" (PI 116). One should always ask oneself in what special circumstances an expression is actually used in order to understand when it makes sense to use it (PI 117).

What is needed instead of theories, hypotheses or explanations, is a *perspicuous representation*. This concept is "of fundamental importance for us" and it "brings about

the understanding which consists precisely in the fact that we “see the connections”.” (RF, p. 69). A perspicuous representation is the result of our investigation of what circumstances licence the use of a word. As such, it is thus never finite since the number of circumstances in which we could use a word such as “red” is infinite. It is however, bound by certain rules that we can find in our use of the concept. These grammatical rules are what we learn when we learn the use of a word, even though we cannot specify them explicitly. It is an overview of such rules that we need to obtain. But this is in fact something most language users already have, since they know perfectly well how to use the word “red”. Wittgenstein says:

[O]ne must only correctly piece together what one knows, without adding anything, and the satisfaction being sought through the explanation follows of itself. (RF, pp. 62-3)

Instead of coming up with an explanation, we can start to see the *connecting links* between concepts/phenomena that are already there in the details of our ordinary language use by investigating particular examples of it. As Wittgenstein says: “A main source of our failure to understand is that we do not *command a clear view* of the use of our words.” (PI 122). Hacker clarifies the notion of a perspicuous representation as follows:

A perspicuous representation is a rearrangement of the rules of grammar that lie open to view, but are not readily taken in as a whole. They become surveyable by such a rearrangement. (Hacker, 2001, p. 341)

When we have reached such a new overview of the use of our language, we will see that “nothing out of the ordinary is involved” (PI 94), and that the essence of the phenomena we are investigating already “lies open to view” (PI 126). It is the role of philosophy to bring about such an ordering or *Übersicht* of a philosophical area in order to clarify the confusions that are being caused by a one-sided view of the phenomena at issue. The ordering philosophy gives us, however, is not *the* order, but only one out of many possibilities that are able to give us complete clarity (PI 132). A grammatical investigation will reveal that everything we need in order to understand our philosophical problems is already there: we only need to see it in a different order. Thus, the result of a good philosophical analysis is the removal of nonsense from philosophy, by showing how we have gone astray by taking the use of our words the wrong way. Then, “the philosophical problems should *completely* disappear” (PI 133). Wittgenstein’s grammatical investigation produces “a kind of understanding which consists in seeing a pattern or form in what is there before our eyes, but which we had previously neglected

or overlooked” (McGinn, 1997, p. 26). All we need to do in order to see the subject from a different, enlightening angle is to rearrange what we already have.

3.5 Therapeutic approach

Wittgenstein describes his own method on several occasions as therapeutic, as a means to treat the philosophical disease of conceptual confusion (cf. PI 119, 133, 254, 593). The way a philosopher should treat a philosophical question is “like the treatment of an illness” (PI 255). And after such philosophical treatment, one should be cured from the conceptual confusions that surround us like a fog (PI 5). This fog clouds our understanding and comes forth from our own preconceived ideas and false pictures. Commanding a clear view of our use of words will disperse the fog and make clear vision possible.

The analogy between a therapy for a disease and Wittgenstein’s philosophical methods highlights three aspects of the latter. Firstly, it shows that the investigation is a slow process in which we eventually come to see how things stand. Step by step the philosophical patient is guided towards a better understanding of the true nature of his problems. Secondly, it reflects the fact that Wittgenstein’s method is aimed at changing our style of thinking. Wittgenstein is not trying to convince us of a certain fact or theory, but he is trying to change our approach towards philosophical problems: he wishes to do away with the theoretical attitude. Thirdly, it shows that our theoretical attitude is *itself* part of the problem: it is only because we adopt this stance towards our philosophical questions that we get into problems in the first place. It is precisely because we try to explain the use of a word, say “red”, by trying to find the essence of that word²¹.

Understanding Wittgenstein’s remarks requires active participation from the reader, who is conducting the investigation together with Wittgenstein himself. The reader is expected to work upon herself in order to see the roots of the philosophical confusions. The fragmentary and repetitive nature of Wittgenstein’s writings are thus *not* the result of an incomplete philosophical vision on the part of Wittgenstein. Rather, they are Wittgenstein’s way to drag us as reader into the investigation with him, in order to actively change our style of thought and attitude towards philosophical questions.

It has to be noted at this point that the therapeutic nature of Wittgenstein’s work has been the subject of recent debate. Recently the book ‘The New Wittgenstein’ has appeared and in it a number of papers is delivered, that according to Crary, share the following viewpoint on Wittgenstein’s work “[Wittgenstein] hopes to get us to see that

²¹ Which, in this case, would probably be something like the frequency-range of red light. However, this will not amount to an explanation of the concept “red” since what we have learned to call red light is dependent on our concept of light and not just on a specific frequency-range.

when we envision ourselves occupying an external point of view on language we don't succeed in articulating *any* thoughts – and that he sees our difficulty as one of coming to recognize that the idea of such a point of view creates the *illusion* of understanding the sentences we want to utter in philosophy.” (Crary & Read, 2001). Moreover, people like Conant and Diamond express in the same volume their views that the main aim of Wittgenstein's work (in this case, the *Tractatus*) is to help us realize that such an external viewpoint is an illusion by going through a number of inconsistent stages within his own work. Afterwards, the reader is expected to realise that the bulk of Wittgenstein's writing is in fact inconsistent or nonsense. However, it feels unlikely that no actual viewpoint or method is to be derived from Wittgenstein's work. I hold that there *is* a consistent reading of Wittgenstein's work that allows for a positive characterisation as well: part of the way Wittgenstein criticizes misconceptions lies in showing how we should approach the subject – in this case, representation – instead. Moreover, Hacker suggests that there is hardly any historical evidence for such a dialectical reading of Wittgenstein (Hacker, 2000). Therefore, I will use Wittgenstein's remarks both to criticize erroneous viewpoints on representation as well as providing an outline of a positive account on representation.

3.6 Methods and Techniques

This final section will provide a short overview of a number of important concepts within the work of Wittgenstein. I will discuss Wittgenstein's notions of *family resemblance*, *language games*, *non-verbal behaviour* and that of *grammar*.

3.6.1 Family resemblance

Wittgenstein's remarks often seem to repeat the same point in different ways, but in fact he is approaching a phenomenon from different directions in to counter the tendency of essentialism (cf. PI 113; 116; section 3.3.1):

A main cause of philosophical disease – a one-sided diet: one nourishes one's thinking with only one kind of example. (PI 593)

The notion of *family resemblance* is crucial to Wittgenstein's criticism of essentialism, for it enables us to get rid of a one-sided viewpoint of the way language functions. In the *Philosophical Investigations*, the concept is introduced in paragraphs 65-67 by considering the different things we call 'games'. If we search for a single common characteristic that binds all sorts of games together – and which would allegedly be the basis for our calling all these things 'games' – we will find that there is no such property. Rather, all sorts of games are related in many different ways, sometimes this property is

common to a number of games, but then another one is common with a different group of games.

Wittgenstein's point is that there is no use searching for a property that is both sufficient and necessary to define the concept of a 'game'. Even if such a defining feature could be found after careful examination, it would not constitute the concept of 'game', for people do not have common characteristic forms as the basis of our calling all these activities 'games': "For if you look at them, you will not see something that is common to *all*, but similarities, relationships, and a whole series of them at that. To repeat: don't think, but look!" ²² (PI 66). Moreover, there are also concepts for which no sufficient and necessary conditions can be found²³, even though we still know perfectly well how to use the concept: think for example about "The ground looked roughly like this" (PI 70) or "Stand roughly there" (PI 71). Given the right circumstances, such expressions can be perfectly clear both to understand and to execute. It is perfectly clear what the correct execution of "Stand roughly there" is. Wittgenstein is thus strongly opposing any form of *essentialism*, which tries to find the essence of words by looking at such shared features.

Instead of finding a common feature, a concept can e.g. be explained by means of examples. Someone giving a number of examples of the application of a word, such as "*this* is red", "an apple is red", "your face is red when you are blushing", etc. is in fact sharing his understanding about a concept. As Wittgenstein says: "Here giving examples is not an *indirect* means of explaining – in default of a better. For any general definition can be misunderstood too. The point is that *this* is how we play the game." (PI 71).

After comparing different sorts of games Wittgenstein concludes: "we see a complicated network of similarities overlapping and criss-crossing" (PI 66). The way various family members resemble each other is similar to this sort of network: sometimes the nose is the same when we compare a given family member to a relative, but when compared to another relative, a different property might be similar. One important concept to which Wittgenstein ascribes family resemblance is 'language': there is nothing common to all that we call language, but all the phenomena "are *related* to one another in many different ways" (PI 65).

²² In PI 67 Wittgenstein introduces the term 'family resemblance' to cover the idea of similarities and relatedness between instances of a concept, whilst moving away from the idea of a common (Platonic) essence.

²³ It could be argued that the common feature of all games is that there are rules involved. Children randomly throwing a ball against a wall can also be said to play a game, but this seems to be a stretching of our normal concept of game. This – I feel – should be called 'playing', but not 'playing a game'. Nonetheless, the fact remains that our use of the concept 'game' is not determined by such a possible common feature.

3.6.2 Language games and form of life

Wittgenstein uses the concept of a *language game* to refer to a number of things. For example, the various activities by means of which we teach children the use of language are part of a language game. If I teach a child colour names, I'm for example pointing to colours and saying the name: "*this* is red", "*this* is green" etc. In order to understand these explanations, a child already needs a certain mastery of language. It must for example be able to play the language game of naming things. The notion language game can also refer to the various practices in which we use language purposefully, for example: giving an order and executing it, reporting an event, telling a joke, describing an object, etc. (PI 23). But Wittgenstein also invents fictional language games as an object of comparison for our normal language use. The notion of a language game is based on an analogy between language and games. This analogy is supposed to bring out two points. First of all, to show that like games, the use of language cannot be defined in terms of one common essence (e.g. 'the description of reality'). Instead of generalising from a number of examples of a use of a word, Wittgenstein consequently points to the differences between such examples. In doing so, Wittgenstein wishes to get away from our need to abstract and generalise. Language games are thus part of Wittgenstein's strategy against the theoretical attitude. It is only when we look at language in isolation – i.e. abstract it from its employment in our everyday lives – that it becomes something problematic. A second strategic role of language games is to show that "language functions within the active, practical lives of speakers" and that "its use is inextricably bound up with the non-linguistic behaviour which constitutes its natural environment" (McGinn, 1997, p. 43). Just like moves in a game of chess have meaning only within the context of chess, moves in a language game have meaning only within the context of the activity in which the language game is played.

Wittgenstein's notion of *form of life* is important to consider here. This idea brings out Wittgenstein's emphasis on language in use, intertwined with the (non-verbal) practices in which we use these words. The concepts language game and form of life serve to show that when we contemplate language, we must look at language in use, not language in isolation. The non-linguistic 'natural environment' of behaviour and practices in which language is employed is fundamental in considering the actual words used: without their 'natural environment' the words become empty as it were. This is why Wittgenstein writes: "Every sign *by itself* seems dead. *What* gives it life? – In use it is alive. Is life breathed into it there? – Or is the *use* its life?" (PI 432). And in PI 38: "philosophical problems arise when language goes *on holiday*". Wittgenstein is here alluding to the fact that by looking at language in isolation we get into philosophical

problems²⁴. In PI 132 he writes accordingly: “The confusions which occupy us arise when language is like an engine idling, not when it is doing work.” It is only when we abstract the normal use of a word that we start to conceive of it as bewildering that words can *mean* anything.

Non-verbal behaviour can count as part of language as well, with clear examples such as pointing, gesturing, waving etc. For these examples the meaning depends on the context just as much as it does for words. The act of waving can mean many things, depending on the context: you could wave someone goodbye, say hi, or even use it to cheer. The notion of a language game serves to bring forward these connections between language and the system of practices that binds a linguistic, social community together (cf. McGinn, 1997, p. 51). Language must be seen as embedded within a historical and cultural community of people that *use* language in their everyday life for various purposes: “to imagine a language means to imagine a form of life” (PI 19). It is when we look at *concrete examples* of how words, expressions, sentences etc. are used in order to do things, that we see the actual characteristics of language that are needed in order to dissolve our philosophical problems. This is why Wittgenstein writes:

We are talking about the spatial and temporal phenomenon of language, not about some non-spatial, non-temporal phantasm. (PI 108)

Looking at the way we teach language to children is one of Wittgenstein’s ways to show how the use of a word is actually being explained. By looking at this process, we are able to get a better understanding of the grammar of a word. Furthermore, Wittgenstein’s fictional language games are devised in order to bring out certain specific aspects of our own language by means of comparison. By showing how, in a very primitive or simple case, certain language utterances would – or would not – work, Wittgenstein is able to draw our attention to particular aspects that we are likely to oversee. In general, if we want to understand the meaning of ‘X’, it will help us to look at how we employ ‘X’ in our language game and to invent simple language games that contain ‘X’ (cf. Z 345)²⁵.

The notion of a language game also brings out that there are many possible ways to employ language. Language games are a family resemblance concept: the diverse functions cannot be reduced to a single option, such as ‘the description of reality’. Wittgenstein gives us a number of possibilities that brings out the multiplicity of

²⁴ For example, such problems as “What is the relation between name and thing named?” (PI 37) or “How do sentences manage to represent?” (PI 435). It is Wittgenstein’s aim to show how these questions are misconceptions, amongst others, by pointing to the fact that language must be considered *in use*, not in isolation. In chapter 4 this will be developed more thoroughly.

²⁵ Note that an investigation into the meaning of ‘X’ is in fact an investigation into the grammar of ‘X’.

language games: giving orders, describing an object, forming a hypothesis, making up a story, making a joke, translating from one language into another, asking, greeting, etc (PI 23; cf. PI 249, 363, 630; Z 345). By no means does this enumeration represent a full account of the possible language games, but it serves to show how many different things we do with language every day. Looking at such actual uses of language will force us to appreciate the diversity of language and will restrain our tendency to theorize about *the* function of language.

3.6.3 Grammar

Wittgenstein applies the notion of *grammar* not only to propositions, but also to expressions, words and phrases (cf. Glock, 1996, p. 150). To Wittgenstein, the grammar of a language is the overall system of rules which determine what does or doesn't make sense to say in the language. If you understand a word, you understand its grammar. So, if we want to explain the use of the word "red" to someone we have to teach her its grammar. A good way to explain Wittgenstein's notion of *grammar* is to contrast grammatical propositions with empirical propositions. A grammatical proposition tells us something about the meaning of a linguistic structure. It gives us information about the rules for using the linguistic structure and thereby shows what kind of meaningful propositions we can form with it. In this way, it tells us "what kind of object anything is" (PI 373): the grammar of 'X' shows us what is to be called 'X'. A grammatical proposition does not state how matters stand in reality, but its being true or not is dependent on the specific pattern of use in our language of the linguistic structure in question. In fact, the specific pattern of use of a certain linguistic structure is a part of reality as well. What is meant here is that the contrast between a grammatical and empirical proposition is that the former expresses a rule of our language (either correctly or not) and the latter represents how things stand in reality (either correctly or not). The former is an explanation of a concept, whereas the latter is a description of a situation. In contrast to this, an empirical proposition does tell us something about how matters stand in reality by using the rules of our language. What grammatical propositions are about is already presupposed in empirical propositions.

An example will serve to shed light on the difference: saying "red is a colour" is a grammatical proposition for it tells us something about how we can apply the word 'red' in our language. It does *not* say that a certain entity, red, is part of another class of entities, objects. Rather, such sentences can help to explain someone the word 'red' if one is unfamiliar with it. Once we have mastered the correct application of this word, we are able to use it in empirical sentences such as "that flower is red", which do state something about the state of affairs in the world (note that this sentence can be either true or false, dependent on the context). As McGinn says: "it is by making ourselves aware of

[the] distinctive patters of use that we clarify the grammar of our concepts” (McGinn, 1997, p. 14).

The notion of grammar does not so much relate to school-grammar or –syntax, as to the meaning of linguistic structures. Most grammatical rules – in Wittgenstein’s sense – cannot be explicitly specified like the grammatical rules we are taught in schools. They are simply part of our ability to use a word. If I say “That flower looks loud to me” or “Can you pass me red please” these sentences do not make sense to us because they do not fit within our grammar of the words used: the words “loud” and “red” can simply not be used in this way, i.e. their grammar does not allow for it²⁶. In PI 11 Wittgenstein compares words to the tools in a tool-box: they all have a different function and need to be applied in a different way. This analogy also reflects the fact that words, like tools, need to have a place within our form of life. We need to learn when and how to apply specific linguistic structures, just as we need to learn when and how to use a hammer.

Wittgenstein makes a distinction between *surface* and *depth grammar*. The surface grammar of “I have a pin” is similar to that of “I have a pain”, but their depth grammar is entirely different: the words have a different pattern of use; the ‘special circumstances’ in which they are applicable differ significantly. Likewise, the ‘to mean’ might look like an activity like ‘to run’ but it is different, meaning is not a process in the same sense as running is a process. Thus, when Wittgenstein says that we are being misled by the form of our language, he means that we are being deceived by the surface grammar similarities. It is by means of his grammatical investigation that Wittgenstein wants to bring the depth grammar of such cases into view: a perspicuous representation will clarify the depth grammar of the considered linguistic structure.

²⁶ Of course, we can always imagine a situation in which such sentences would make sense. However, the point is that words can simply not be used in every possible way: some words require a different context and application than others.

"The aim of philosophy is to erect a wall at the point where language stops anyway."

- Ludwig Wittgenstein

Chapter 4

A Critique of the Causal Theory of Representation

4.1 Introduction

In Chapter 2 we have seen that the subject of representation poses an intriguing and challenging problem that is still very much alive in philosophy. It is a problem that has kept philosophers busy for a long time and will most likely remain attracting philosophical interest. Even apart from the biological and cultural evolution of a system of representation like language, the human ability to *use* representations in our daily practices is something wonderful indeed: it allows for knowledge and experiences to be passed on generation after generation. Obviously, a facet so pervasive in our lives can expect to receive the attention from both philosophers and scientists. Then again, is language really something unique or mystical? Wittgenstein writes the following:

“Language (or thought) is something unique” – this proves to be a superstition (not a mistake!), itself produced by grammatical illusions. And now the impressiveness retreats to these illusions, to the problems. (PI 110)

At the same time, Wittgenstein also urges us to see the particular details of each use of language. It is because we are focussing at the wrong ‘problems’ that we do not arrive at the understanding of what representations are. The problems we are facing here are only apparent problems, arising from our misunderstanding of the forms of our language:

They are deep disquietudes; their roots are as deep in us as the forms of our language and their significance is as great as the importance of our language. –

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

Let us ask ourselves: why do we feel a grammatical joke to be deep? (And that is what the depth of philosophy is). (PI 111)

We see that Wittgenstein alludes to a distinction between the type of difficulty faced in philosophy, and that faced in e.g. the natural sciences. It is by a better understanding of our language – i.e. by understanding why a grammatical joke has the character of depth – that we will be able to solve the philosophical puzzles at hand, not by making empirical discoveries or experiments.

This chapter envisages a critique on the causal theory of representation as outlined in Chapter 2. It aims to disentangle the misunderstandings present in the views of both Dennett and Dretske, as representatives of the causal theory of representation. With the help of Wittgenstein's method as described in Chapter 3 and several of his remarks I will critically investigate the viewpoints of Dennett and Dretske. Since Wittgenstein's remarks are vastly interconnected and the issues involved were considered by him on various occasions²⁷, I will draw from various sources, such as the *Philosophical Investigations* and the manuscripts "Cause and Effect"²⁸ and "Remarks on Fraser's Golden Bough"²⁹.

4.1.1 Chapter Outline

The first part of this chapter will focus on the differences in methodology between Wittgenstein on the one hand and Dennett and Dretske on the other. I will claim that both Dennett and Dretske fail to make a proper distinction between causes and reasons at the start of their investigations. Wittgenstein brings out this difference to show, amongst other things, the fundamental discrepancy between empirical science and philosophy (i.e. what he thinks philosophy should be). Not appreciating this difference results in a misunderstanding: trying to apply the hypothetico-deductive method of the natural sciences to the problem of representation.

I will then go on to show that neglecting this difference is also one of the roots of Dennett's mistaken reception of folk-psychology as a (scientific) theory or a semi-scientific model for interpretation. In spite of what Dennett believes, folk-psychology is not an attempt at science, i.e. it is not a causal explanation. Rather, it is a rational

²⁷ Cf. PI, preface: "I have written down all these thoughts as *remarks*, short paragraphs, of which there is sometimes a fairly long chain about the same subject, while I sometimes make a sudden change, jumping from one topic to another." This shows that even within the *Philosophical Investigations* Wittgenstein moved back and forth between subjects.

²⁸ Referred to as CE, printed in: *Human Knowledge, Classical and Contemporary Approaches*, Paul K Moser & Arnold vander Nat (eds.), Oxford: Oxford University Press.

²⁹ Referred to as RF, printed in C. Luckhardt (ed.), *Wittgenstein. Sources and Perspectives*, pp 61–81, Ithaca, N.Y.: Cornell University Press.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

explanation³⁰ of why people do what they do, an explanation that gives meaning to our lives. In addition, Dennett's idea of the intentional stance – fundamentally related to his conception of folk-psychology as a theory – is corrupted by problems concerning interpretation: are we really willing to say that *all* mutual human understanding is the result of a successful interpretation? Wittgenstein opposes this and with the use of his remarks I will show that there are also ways of understanding people's behaviour – including verbal behaviour – that are *not* an interpretation. I will furthermore discuss two other problems concerning the idea of successful communication as the result of interpretation, i.e. the problem of 'mentalese' and a problem concerning the understanding of intentionality.

Whereas the former problems are most explicit in Dennett's work – although they play a role in Dretske's work as well – the next problem is more explicit in the view proposed by Dretske. As stated before, both Dennett and Dretske see the relation between a representation and what it represents as an external relation. In Dretske's work this relation takes the form of a nomic indication (Dennett subscribes to a similar viewpoint, cf. Dennett, 1987, pp. 30-31). But as an explanation of representation this proposal faces a problem already introduced in Plato's *Theaeteteus*, namely that of misrepresentation: how can a representation represent something that is not there/not the case? Dennett and Dretske both provide similar solutions (talking about frogs), but in my view these fail to solve the problem. However, there are more serious difficulties with the indication approach: Both Dennett and Dretske explain representation as an external relation between two independent events or entities. In the 1930's Wittgenstein already criticized others on this point: the 'relation' between a word and its meaning is not external, but rather internal. In order to show this I will introduce the difference between a sign and a symptom, showing that e.g. 'clouds are a sign of rain' is a whole different matter than "red" is our sign for red'.

Finally, I will go on showing how Wittgenstein positively conceives of the relation between representation and what is represented, or between word and object, by focussing on learning, training, language games, form of life and particular examples.

³⁰ I've chosen to use the concept "rational explanation" because it conveys the idea that reasons are involved. This type of explanation is sometimes also labelled "reason-giving explanation", but I find that label less appealing. Jeanne Peijnenburg has used the terms "action explanation" and "reason explanation" for this type of explanation (Peijnenburg, 1998, p. 81). Note that I do not intend to distinguish deliberate, planned behaviour that is the result from a process of ratiocination from subconscious decisions on this point. I simply wish to distinguish between explanations in terms of reasons and explanations in terms of causes.

4.2 Methodology

So far, we've seen a number of differences between Wittgenstein and Dennett and Dretske. Whereas the latter support the hypothetico-deductive method of the natural sciences, forming hypotheses and trying to verify these against experimental data, Wittgenstein follows his own method of the grammatical investigation. A second difference is that where Dennett and Dretske aim at finding and composing causal explanations, Wittgenstein gives conceptual explanations of the problems discussed. In the subsequent section we will go into these two differences between Wittgenstein on the one hand and Dennett and Dretske on the other.

4.2.1 Philosophy deals with a different category of problems

In the previous chapter about Wittgenstein's method of philosophy we discussed how and why Wittgenstein believes philosophy is not like (empirical) sciences such as physics or chemistry. We've also seen that both Dennett and Dretske adopt more of a scientific approach towards the problem of representation. As noted in Chapter 2, Dennett explicitly declares the objective, materialistic and third-person viewpoint of the physical sciences as his starting point (Dennett, 1987, p. 5) and aims to find a scientific explanation in terms of causes and their effects as to how a physical structure like the brain can accomplish something like discriminating meanings (cf. Dennett, 1994, pp. 136-7). Dretske has a similar point of view: he feels that there *must* be some kind of explanation as to how a purely physical system can occupy states that have a meaning, i.e. states that are representations (Dretske, 1980, p. 355). Like Dennett, Dretske wants to make sense of representation within the framework of the causal world picture, which results in his approaching the problem in a scientific manner.

Wittgenstein sees the problems of philosophy as being different in nature and consequently believes that they are in need of a different approach: his method of grammatical investigation leading towards clarification of concepts instead of a theory or model that tries to explain the (causal) mechanisms at work. Scientific enquiry, i.e. the collection of data and the testing of hypotheses, is of no use when we are confronted with a conceptual instead of an empirical problem. Hillary Putnam, who has also been concerned with the problem of representation throughout his philosophical career, writes: "Analytic philosophy has become increasingly dominated by the idea that science, and only science, describes the world as it is in itself, independent of perspective" (Putnam, 1992, pp. ix-x)³¹. Like Wittgenstein, he seems to think that this idea is getting too much

³¹ Interestingly enough, in this book (*Renewing Philosophy*) Putnam attacks his old views, e.g. as known from *The Meaning of Meaning* (1975). In his later work he seems to have a much more

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

grip within philosophy. Shortly afterwards he adds: “Nevertheless, the idea that science leaves no room for an independent philosophical enterprise has reached the point at which leading practitioners sometimes suggest that all that is left for philosophy is try to anticipate what the presumed scientific solutions to all metaphysical problems will eventually look like.” (Putnam, 1992, p. x). Again we can compare this with Wittgenstein’s remarks about the role of philosophy and the nature of philosophical problems (see Chapter 3). Clearly, both philosophers think that philosophical problems cannot be reduced to the level of physics, and therefore cannot be answered within the causal framework, i.e. in a scientific manner.

Note that Wittgenstein does not question the causal world picture in general, but simply disagrees that it has anything to offer for our problem (i.e. the problem of representation). Part of the mistake lies in presupposing the causal world picture as *the* proper background for the problem of representation within which we should solve it. In order to solve our problem, i.e. in order to climb out of the philosophical darkness of the pit, we should conduct a grammatical investigation aimed at providing a perspicuous representation of the relevant concepts. No experiments can shed light on this issue, only a better understanding of the various concepts at hand.

Likewise, I do not wish to quarrel with proponents of the causal world picture. Indeed, the brain is a physical structure and is as such subject to the causal laws of nature as is any other physical thing. And indeed, we do have nerve cells that send electrical pulses when triggered by certain stimuli, and these electrical pulses travel to the brain, where in turn other nerve cells are activated which results in muscle contractions and the like. All these things are physical and obey the laws of nature. This is beyond doubt. At this point it is interesting to evaluate a description given by Dennett of a frog’s escape behaviour:

As it lies poised motionless on the lily pad, its nervous system is humming with intricate activity, the products of millions of interactions among photons, acoustic pressure waves, receptor cells, internal secretions, and the like interact with each other to produce yet more activities, which eventually yield among their sums the efferent pulses that contract the frog’s leg muscles and send it hurtling leftward into the net. (Dennett, 1987, pp. 108-109)

This description is a perfect example of the causal processes at work when the frog is exhibiting escape behaviour. But the point is: *this* is not what we call escape behaviour.

Wittgensteinian approach towards philosophical problems. Related to this he seems to think that philosophy and science do differ fundamentally and he disagrees with the reductionistic framework aimed at by many contemporary philosophers. Compare this to Wittgenstein’s view on the role and aim of philosophy described in Chapter 3.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

We know perfectly well what escape behaviour is without being able to specify all these things about the frog's nervous system and muscles. What we have learned to call escape behaviour is something different, something which we compare to the paradigm of human behaviour, and which can be applied to other animals as well. That is, it is a family resemblance concept: there is no essence we can find that makes a certain chain of events escape behaviour and other chains not³². Dennett's description is accurate and useful (to e.g. neuroscientists), but it is not the realm where we should look if we wish to understand the frogs representations. It is the place where we should look if we want to understand the physical workings of the brain, nerve cells, electrical signals, neurotransmitters and so on. Having a good model of the brain does not mean that one understands what a person will decide, nor that one will understand why someone has made a certain decision. It will only provide us with information about – and possibly an explanation of – the workings of the brain itself – which, admittedly, can be a very interesting and fruitful enterprise in itself. But it will not tell us something about the persons mind.

Talking about representation is, despite what Dretske and Dennett claim, not the same as talking about physical systems. They feel that the biggest problem of representation is the question of how a purely physical/mechanical system can be said to have a representation, i.e. how a purely physical system can exhibit aboutness. However, what we can learn from Wittgenstein is that this only seems mysterious because we try to explain representation in such a mechanistic way. But causal explanations of the mechanisms of our nerve cells or information bearing structures will *not* help us to solve our philosophical puzzle of representation. The problem of representation as Dretske and Dennett see it, only springs into view when we try to fit 'meaning' into a causal view. But the concept of meaning does not fit into the causal scheme.

At this point it might almost seem as if the concept of a representation becomes superfluous. But that is not the point of the argument. There is no denying of the existence of representation, just the claim that a causal explanation of it is nonsensical. To understand intentionality we have to adopt a different stance: one that does not look for an explanation, but one that searches for a clarification, i.e. one that will in the end resolve us from the urge to explain representation within the causal world picture.

In conclusion we can say that the method supported by Dennett and Dretske is fundamentally different from that used by Wittgenstein. This is related to what both

³² Compare this to what Dretske writes about a cat's stalking behaviour: "To stalk another animal is not merely to retain spatial proximity to it, but to do so with a certain lethal purpose. If a cat's movements do not have a special kind of etiology, then the cat isn't *stalking* a mouse" (Dretske, 1988, p. 8). This is a perfect example of the essentialism Wittgenstein is battling against, as Dretske is here trying to analyse the essence of staking behaviour, which has to do with a particular origin, or history, of the behaviour.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

parties see as the real problem (and related to this, the type of solution needed): whereas Dennett and Dretske think we have to specify the underlying mechanism of a certain phenomenon, Wittgenstein thinks that we need to get a different understanding of representation in the first place. This is reflected in their viewpoint on the nature of the problems: Dennett and Dretske are convinced they have a causal, empirical problem at hand while Wittgenstein is convinced that the problem is conceptual: it is the result of misunderstanding the forms of our language. The difference between problems that can be solved with the use of the hypothetico-deductive method and the conceptual problems Wittgenstein is aiming to solve, is connected with the difference between causes and reasons. This important will be clarified by the subsequent discussion. A final difference is that we can see Dennett and Dretske generalising and abstracting from a number of examples to construct a general theory as an explanation for representation. Wittgenstein on the other hand directs our attention towards specific examples of language use, thereby showing us differences between cases. Whereas Dennett and Dretske show signs of essentialism, Wittgenstein is battling this, e.g. with the use of his family resemblance concept.

4.3 Causes versus Reasons

The first issue we will be looking into is the difference between causes and reasons. I believe that this difference can shed light on the theories proposed by Dennett and Dretske, and can be used to criticize the causal theory of representation. As we shall see, the causal world picture in itself is not problematic – something which is supported by the enormous progress and indispensable solutions that have been achieved within this framework in the empirical sciences – but the question is whether it is the appropriate framework for when we are concerned with questions of meaning and representation.

4.3.1 Dennett and Dretske on causes and reasons

Dennett writes that beliefs are information-bearing states that arise from perceptions and lead to action (Dennett, 1987, p. 46). Mental states, i.e. representations, are thus said to have causal powers. He also writes, considering the escaping frog again, that we interpret his behaviour as “wanting to escape”: in the end the frog’s eyes contribute to the control of the frog’s limbs, and *that* is what we call ‘beliefs’ (Dennett, 1987, p. 107). Shortly after, Dennett writes:

... what reason is there to suppose that human belief is all that different from frog belief? In both cases behaviour is controlled by a complex internal state

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

that can be alluded to more or less effectively by the everyday folk practices of belief attribution and desire attribution. (Dennett, 1987, p. 115)

Dennett is saying here that the “complex internal states” are in fact what we normally call beliefs. In this way he in fact attributes causal powers to mental states. Likewise, Dretske writes:

Experiences and beliefs are merely those internal, presumably physical, states of a system having the function of providing information (in the case of experience) and mobilizing it (in the case of belief) for use in the control of behaviour. (Dretske, 1994, p. 260).

Here we see that Dretske also presents mental states as being identical to system states that in the end control behaviour. The way Dretske describes the functionality of these states is another example of a causal explanation: for example, walking to the kitchen to get a glass of water can simply be described as the (physical) result of a certain system state that, most likely, has something to do with dehydration and which has accordingly activated other internal states and processes that in the end result in you walking to the kitchen. Dretske repeatedly directs us to the idea that reasons are nothing more than internal causes: in the preface to *Explaining Behaviour* he writes “My reasons, my beliefs, desires, purposes, and intentions, *are* – indeed they must be – the cause of my body’s movements” (Dretske, 1988, p. ix). In chapter 1 of the same book we find Dretske to make a distinction between external and internal causes: whereas stones are completely subject to external forces (gravity, temperature, etc) conscious beings can behave in different ways as the result of their having internal causes as well (cf. e.g. Dretske, 1988, p. 7). Dretske tells us that *behaviour* is in fact “the production of external effects by internal causes” (Dretske, 1988, p. 33), hereby depriving cases such as “the behaviour of a falling stone” from the application of the concept behaviour. He distinguishes between behaviour and movement in the following way: if a rat moves his paw, the movement of the paw is just a movement, but the rat moving its paw is behaviour. A falling stone would thus fall under “movement” rather than “behaviour”. For the latter includes the bringing about of the movement by internal causes, according to Dretske, which is not the case here.

Excluding a falling stone from the range of application of “behaviour” is odd, because people use expressions like “the behaviour of a falling stone” regularly and understand perfectly well what is meant by it. (Indeed, one could also say “the movement of a falling stone” but the fact remains that the other expression is also used commonly). But this is not because the behaviour of a falling stone has some kind of essential feature in common with the behaviour of, say, someone getting a glass of water, but because

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

‘behaviour’ is a family resemblance concept that we have all learned to apply and understand correctly³³.

In conclusion we can say that both Dennett and Dretske do not see reasons and causes as two exclusive categories, but rather classify the former as a special (i.e. internal to an intentional system) case of the latter. Thus, we see Dennett writing on e.g. the evolution of life that “self-replicating macromolecules *had* reasons for what they did, but had no inkling of them” (Dennett, 1987, p. 64). And later, on the behaviour of a fledging cuckoo “[t]he newborn cuckoo is, of course, oblivious; it has no inkling of this rationale for its ruthless act, but the rationale is *there*, and has undoubtedly shaped this innate behaviour over the eons.” (Dennett, 1987, p. 65).

4.3.2 The difference between causes and reasons

Let us now try to see why causes and reasons should be considered as fundamentally different³⁴. The distinction between causes and reasons can be shown by considering the type of explanation they are used in. Two issues are at play here. First off, Dennett and Dretske take what are rational explanations of (human) behaviour as if they are causal explanations. Secondly, the explanations that Dennett and Dretske offer for behaviour will not satisfy someone who is after the *reason* of the behaviour. We will first look into the difference between rational and causal explanations and then take on the second issue.

Dennett and Dretske claim that mental states are causes to behaviour: they are seen as internal processes that have gained causal power over external/output behaviour (e.g. Dennett, 1987, p. 31; Dretske, 1988, p. ix). However, Dennett and Dretske are being

³³ One might even say that “behaviour” when applied to inanimate objects is a secondary sense of the word that is related to, but not the same as, “behaviour” applied to animate beings. For more discussion on secondary sense see Ter Hark (1990).

³⁴ Jeanne Peijnenburg has suggested, following Davidson, that causality has a dual character in order to bring reasons and causes closer together (Peijnenburg, 1998). Instead of following either the causalists, or the adherents of the Logical Connection Argument (LCA) – an approach that comes close to Wittgenstein’s viewpoints on the difference between causes and reasons – she claims that Davidson has found an in between solution. According to Peijnenburg, Davidson thinks that rational explanation is explanation in a broad sense, while causal explanation is explanation in a strict sense. Without going into the details here, one could question whether it is justified to combine two different types of explanation under one header: isn’t that in fact an argument against Davidson’s combination? I will not go into the details of the solution here, but Davidson’s solution to combine reasons and causes under one header seems to amount to the following: he believes that both causal and rational explanation of actions are relevant, but that they count as explanations in a different way. I think however, that the view adhered to by Wittgenstein (and myself) is not suitable for such an integration. The difference between causes and reasons is too fundamental to be combined with one another. See Peijnenburg (1998) for a more elaborate explanation of Davidson’s theory. See also Thornton (1998), chapter 6.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

misled by the surface grammar similarities between certain expressions, even though they function entirely different as explanations. For example, saying “because I was angry at him” is something entirely different from saying “because my arm muscles contracted” when we have just hit someone in the face. Both are similar in surface grammar, but they serve a different explanatory purpose: one is rational, and the other is causal (or physical, if you will).

Freud’s conception of human behaviour provides a similar mistake, so it will be worthwhile to look into his ideas shortly. Like Dennett and Dretske, Freud was looking for an explanation for human behaviour in terms of causal processes. In his theories, Freud generalizes cases of correctly relating a reason to an action by proposing that particular explanation as *the* general explanation. Take Freud’s view that all cases of anxiety are caused by the birth trauma, or the idea that a wooden table in a dream represents a naked woman. Such generalisations are wrong, because they pass over the characteristics and context of a specific case. As Wittgenstein admits, relating anxiety to the birth trauma has a marked attraction (LC, p.43), but this relation has not been established by means of numerous investigations of cases of anxiety. There is no proof whatsoever that it holds for all cases. In addition, Wittgenstein points out in an analogous case of hitting someone, that if “some activity is shown to be carried out often for a certain purpose [...] then a hundred to one it is also carried out under other circumstances *not* for that purpose” (LC, p.44). One case of a correct explanation does, however, not justify us in assuming that all cases can be explained in the same way. Moreover, the purpose of an action is something different than its cause. Generally, we describe the actions of other humans in a rational way, whereas we describe physical events in a causal manner: one time I might hit someone in order to inflict pain, the other time to surprise him or even to praise him. In each case we can ask the person in question “Why did you hit him?” and the answer would look like “Because of such-and-such reasons”. But these surface similarities should not mislead us into thinking that each and every case of hitting can be explained this way. Sometimes we might really be interested in the causal explanation, e.g. when the hitting was done by accident or on purpose. The difference is that a rational explanation tries to give meaning to events, whereas a causal explanation gives us information about how something came to be in terms of the underlying processes or mechanisms.

We clarify the difference between the notions of reason and cause by considering the various uses of *cause* as brought forward by Wittgenstein. In *Lectures and Conversations*, Wittgenstein says:

‘Cause’ is used in very many different ways, e.g.

(1) “What is the cause of unemployment?” “What is the cause of this expression?”

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

- (2) “What was the cause of your jumping?” “That noise”
- (3) “What was the cause of that wheel going round?” You trace a mechanism.
(LC, p.13)

Wittgenstein aim here is to show that there are different uses of the word ‘cause’. Not all these uses can be explained by assuming one and the same *essence* in each occurrence. Even though the three mentioned uses of ‘cause’ may be related, there need not to be one essence that can explain all these uses. Note that I (will) use the concept cause in a more strict manner: Wittgenstein shows that in some occasions one can call a reason a cause, e.g. in the second example here. He distinguishes the three cases later on as follows: “(1) Experiment and statistics. (2) Reason. (3) Mechanism.” (LC, p. 15). For reasons of clarity I will therefore not use the word “cause” in that manner myself. The point here is that the second use of the concept cause is very different from the first and third use of the concept.

An important difference we can establish by looking closer at the concepts of cause and reason, is that causes can be discovered by conducting a number of experiments and drawing conclusions from those, while reasons cannot be discovered in such a manner. Rather, reasons have to be accepted by others (or by yourself) to be taken in as reasons. Jacques Bouveresse, who has written on the subject of causes and reasons against the background of Wittgenstein’s remarks on Freud, describes the difference between a hypothesis (cause) and a reason as follows: “[a] reason is characterized by the capacity to be recognized as such by the person whose reason it is, and not on the basis of an inductive inference.” (Bouveresse, 1995, p.69). Note also that there are differences between what other people will generally accept as (good) reasons, and what they will accept as causes. Take for example the fact that a lot of people are inclined to accept Freud’s explanations. The reason for this inclination is that Freud’s explanations provide the same kind of explanatory satisfaction as mythology: they enable one to make sense of actions and endow them with meaning. For example, Freud’s interpretations of a dream fits it “into a context in which it ceases to be puzzling” (LC, p.45) and “after the analysis of it, the dream appears so very logical” (LC, p.51). Often people will have to be convinced by their psychotherapist that *this* is the correct analysis of their dream or anxiety. This shows a difference with causes: to redescribe one’s actions as being ‘caused’ by a certain motive is not the result of a discovery – i.e. of the motive – but the result of persuasion: “In a different way you could have been persuaded of something different” (LC, p.27). That is: causes can be discovered but reasons have to be accepted. Even in cases where psycho-analysis ‘reveals’ a hidden motive/reason, it is the patients *acceptance* that makes it the reason for his behaviour. Without the patient accepting it, it will not be able to count as a reason. There is no amount of evidence or testing that can

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

support it³⁵. (Note that this is different for causes: if gravity is the cause of a certain process, it will be the cause even if a person seeing the process does not believe in gravity.)

In this sense, a cause is always something hypothetical that one can test, whereas a reason is not. Note that we can only speak of reasons when we consider human behaviour, not when we consider physical events: it does not make sense to attribute reasons to inanimate objects (as Dennett does). When people adopt a rational explanation, the actions or experiences that were problematic or strange before now seem to fit in: people will find it easier to accept them. They are now able to see these actions or experiences in a different light and will describe them in a different way, i.e. they have endowed them with meaning. The problem lies in assuming that reasons are to behaviour exactly what causes are to effects (e.g. forces such as gravity to physical behaviour). Bouveresse describes this difference strikingly in the following way:

... the relation between a reason and the action it explains is a logical and internal relation, since a reason consists of redescribing the action with the effect of making it intelligible, whereas the relation of cause and effect is an empirical and external relation between two events. (Bouveresse, 1995, p.75)

This is a remark about the conceptual status of reasons and causes, meaning that the effect of a cause cannot be placed on a par with the behaviour that ‘follows’ from a reason. It is what Wittgenstein calls a grammatical remark because it shows a difference in use between the terms cause and reason, that is, it shows that reasons and causes belong to different context. It is wrong to think that behaviour results from reasons in the same way as falling is the result of gravity, or as high blood-pressure is the result of a bad diet, because reasons are used in a different manner and context than causes are being used.

The second issue is that the explanations offered by Dennett and Dretske will not satisfy us when we want to know the *reason* of someone’s behaviour. For example, stating the cause of the motion of your limbs does not suffice as an explanation to the question “why did you hit him in the face?”, for the causal explanation provides no answer to the rationale behind the hitting³⁶. The language game we are playing asks for a

³⁵ Note that Wittgenstein says that the difference in grammar between cause and reason is similar to that between cause and motive (BB, p.15). Note also that Wittgenstein points out that in a law-court one is supposed to know the motive of one’s action, whereas one is not supposed to know the physical laws that govern the movements of your body (LC, p.21).

³⁶ Note that Dennett makes similar claims when he speaks of different levels of description (Dennett, 1991b). The problem however is that Dennett still wants to hold onto an ontological connection between the levels: the physical level is what enables the design or intentional level. Whereas Dennett

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

reason-giving explanation, not for a causal explanation. On Wittgenstein's view, there is an important difference in type of explanation: the causal theory of representation tries to offer an explanation in terms of physiological and mechanistic processes in the brain, the body and its environment – that is: a causal explanation. Wittgenstein discusses a similar issue in Part II of the *Philosophical Investigations* when he investigates the phenomenon of aspect perception³⁷. Here we also see the craving of many philosophers to come up with some kind of causal explanation, whereas what is really needed is clarification of the involved concepts. As Wittgenstein writes here, whenever a physiological explanation is offered, "[t]he psychological concept hangs out of reach of this explanation" (PI, p.212). As said, a physiological or mechanistic explanation will be about *what goes on* in the body or brain when someone is, say, uttering a sentence with meaning. But such an explanation will not help us if we want to know a person's reasons for his actions. Such explanations are helpful if we want to understand the workings of his brain and body, e.g. in situations where someone is getting surgery, taking psychotherapy or is being investigated by a psychiatrist. But there is not one type of explanation that will work in each occasion, as the type of explanation needed is dependent on the context. But in the case of aspect perception, such an explanation will not help us to understand the phenomenon any better.

Moreover, having a thought, belief or a representation does not amount to experiencing a certain brain-state. When we have a representation of something, or when we understand what some representation represents, we do not experience our brain-states but we e.g. *understand* or *think of* something. Thinking of something is not a brain-state, it is a type of behaviour that we can all recognise: consider e.g. the situations in which you would say of a person that she is thinking about something. Knowing what kind of neurological states – i.e. nerve cells firing, chemicals reacting, etc – we have during such processes involving representations, will not clarify what representations are or how they function, even though it will give us important insights in the nature and workings of the human brain. Likewise, knowing what goes on in someone's body when he hits someone else in the face is not the same as knowing why he hits that other person in the face. It is this type of difference that Wittgenstein alludes to in the following passage:

The feeling of an unbridgeable gulf between consciousness and brain-process:
how does it come about that this does not come into the considerations of our

places his 'levels' on top of one another, Wittgenstein would rather place the two separately beside one another, without any ontological connections: rational explanation is simply a different language game than causal explanation.

³⁷ This phenomenon concerns the 'switch' between perceptions one makes when seeing ambiguous figures such as the duck-rabbit picture or the Necker-cube.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

ordinary life? This idea of a difference in kind is accompanied by a slight giddiness, – which occurs when we are performing a piece of logical sleight-of-hand. (The same giddiness attacks us when we think of certain theorems in set theory.) When does this feeling occur in the present case? It is when I, for example, turn my attention in a particular way on to my own consciousness, and astonished, say to myself: THIS is supposed to be produced by a process in the brain! (PI 412)

Here we see the difference between Wittgenstein and Dennett and Dretske expressed in a very clear form. Whereas Dennett and Dretske want to find out how it is possible that what you think, feel or believe is produced by a process in the brain, Wittgenstein points out to us that this is an illegitimate question as the difference between brain processes (neurological processes in terms of physics) and consciousness is not simply a matter of degree, but is a difference between language games. That is, it is a difference between the contexts in which we are allowed to use these concepts: talking about brain processes takes place in a different context from talking about reasons, beliefs, or desires. Taking these two to be part of one and the same realm is exactly the “logical sleight-of-hand” we should stay away from – it is the type of mistake that we will not be inclined to make anymore if we have obtained a perspicuous representation.

If we want to know what representation is, we will have to investigate the use of concepts like ‘meaning’, ‘reference’, ‘representation’ and the like. It will not do any good to scrutinize our brains at the moment of having or using a representation. What could even count as ‘the moment of having a representation’? How would we know what kind of brain states would be candidates for “brain-states that exhibit aboutness” (as Dennett puts it)? How would we know when the representation starts and when it ends? Such questions *seem* legitimate but they are based on misunderstandings: representations are not parts of our brains, they are not activation patterns and our brains do not represent, we do.

After having taken in the difference between a rational and a causal explanation, we can now understand why Dennett’s conception of folk-psychology as a naïve form of science is mistaken: folk-psychology never has had the ambition of being a science and the explanations offered in folk-psychology are not meant as causal but as rational explanations. They are supposed to make people have an understanding of the behaviour explained, not in terms of what kind of mechanism is behind it but in terms of its rationale. Folk-psychology is what makes us understand someone’s behaviour in terms of the reasons she provides for it. Such rational explanations endow the behaviour with meaning and therefore have an important role in human live.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

4.3.3 Understanding is not a form of interpretation

However, there is another problem, albeit related, with folk-psychology. Conceiving of folk-psychology as a naïve scientific model for explanation and prediction of human behaviour is a misconception because it presupposes that when we are interacting with other human being we are constantly interpreting. Taking in the rational explanations of folk-psychology as if they are causal explanations and making sense of the behaviour of ‘intentional systems’ in that way is a form of (elaborate) interpretation to Dennett, where folk-psychology offers the key for translation: scratching one’s nose means nervousness, yawning means being bored, asking for water means being thirsty, etc. This way we can predict what people will do next, what they want, what they mean, etc. As Dennett writes: “We use folk-psychology – interpretation of each other as believers, wanters, internders, and the like – to predict what people will do next” (Dennett, 1991b, p. 29). Patterns in the behaviour of intentional systems are signals about their mental states: “The success of folk-psychological prediction, like the success of any prediction, depends on there being some order or pattern in the world to exploit.” (Dennett, 1991b, p. 30). Shortly after this he adds: “the pattern is discernible in agents’ (observable) behaviour when we subject it to “radical interpretation” (Davidson) “from the intentional stance” (Dennett)”. Dennett is opting for what has been called a mild-realism about such patterns, and believes that using folk-psychological concept is a way of dealing with these patterns that is much more efficient and useful than a possibly more reliable system that takes in all the neurological and environmental factors (i.e. a system that describes and calculates what other beings will do on the level of physics, not on the level of psychology). In the end it comes down to this: we make observations about other people, take these in, and with the help of folk-psychological models and concepts, we make sense of what others are doing and can accordingly predict/explain their behaviour.

One problem with seeing prediction and understanding of others as a result of interpretation seems to be the following: if behavioural patterns – amongst which are linguistic ones – need to be interpreted before we can understand them, this presupposes something like *mentalese*³⁸. That is, if we in fact translate what others do with the use of folk-psychology, there surely must be something *into* which we translate it. Thus, we end up needing yet another (private) language into which we translate what we observe, a new language that actually makes sense to us. Obviously, this does not explain the problem of representation, nor that of how we are able to understand other conscious beings: it only shifts the burden of proof one level backwards. In short, Dennett’s ideas commit him to interpretation for each time we understand something, and therefore he

³⁸ The term “mentalese” stems from Fodor, who proposed that there might be a special language of thought (LOT) in which our thoughts are represented. This LOT would be independent from our actual everyday language. See Fodor (1975) for his own explanation of this theory.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

ends up with yet another problem: how is a certain behavioural pattern interpreted if not in a language?

Moreover, Dennett's theory of the intentional stance is in fact claiming that we are constantly busy interpreting other people's behaviour and utterances. It claims that we generally hold a belief or a hypothesis confirmed by experimentation that other beings are conscious, like ourselves. The confirmation is that when we view them as being conscious, i.e. when we use the intentional stance, we can easily predict and understand their behaviour, Dennett claims.

However, what seems to be an inconsistent point in Dennett's philosophy is that on the one hand he defines intentional systems as those systems that are successfully interpreted or explained by the intentional stance, but on the other hand, he defines the intentional stance as a way of interpreting the behaviour of (intentional) systems. What is problematic here, is both that this seems to be circular and that there is no way to learn what intentionality really is except by *assuming* that something is intentional. But in order to assume that something is intentional, shouldn't you already have an idea of what intentionality is? Can one really have a viewpoint towards something *as if* it has such a property, without actually ever having encountered that property without doubt elsewhere? Compare this, for example, with looking at a person as if he would be a robot: can one do this when you have no acquaintance or understanding of robots in the first place?³⁹ Marc Slors has written on this point recently and his main point is similar to what I'm suggesting here:

- (i) [T]he intentional stance theory suggests that mastery of intentional-state concepts depends on one's capacity to adopt the intentional stance. Therefore what it is to adopt the intentional stance must be explained, if at all, without mentioning beliefs, desires, etc., if circularity is to be avoided. (ii) No such non-circular explanation can be provided, given the types of system Dennett wishes to include in the class of intentional systems. (Slors, 1996, p. 93).

Conceptually, Dennett gets into trouble because he either assumes that mastery of intentional-state concepts precedes our ability to adopt the intentional stance, or vice versa. But both options end up in a circularity problem. For Wittgenstein, both the mastery of intentional-state concepts *and* the ability to adopt an attitude towards a soul are things that come in at the same time, in the same way. This is what we learn when we grow up and start to interact with our environment. By learning intentional-state concepts

³⁹ Cf. "“I believe that he is suffering.” – Do I also *believe* that he isn't an automaton? It would go against the grain to use the word in both connections" (PI, p. 178). The word Wittgenstein is referring to is "believe": believing that someone is suffering is something fundamentally different from believing that someone is not a robot.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

we are learning to adopt an ‘intentional stance’ – these points are bound up into one. Correctly understanding the behaviour and language of others is not a matter of hitting on the right interpretation, but it is a matter of having grasped the meaning by means of extensive training and mastering of a practice of using linguistic ‘tools’: it is learning the technique, the skill, to employ language. For example: what “length” means is in part determined by what it is to determine the length of a rod, and the mastery of this latter technique is needed for a correct understanding of “length”. The same holds for folk-psychological concepts, such as “pain”, “thought”, “belief”, “desire” and the like. All these words function within different practices that one masters during ones life. Understanding these words means that you have mastered the specific practices in which it is employed; it means that you have been successfully trained in your form of life.

Wittgenstein would thus object that it is assumed in our everyday actions that others are ‘intentional systems’. An assumption has the status of a hypothesis: one could be proven wrong about it. But that others are ‘intentional systems’, i.e. that they are conscious beings, is a certainty that is not accepted on rational grounds, but forms the basis for rationality itself. Taking the behaviour – linguistic or other – of other human beings is not a hypothesis or a conclusion, no matter how fast or ‘unconscious’ this conclusion is supposedly drawn: it is a certainty in our way of dealing with others⁴⁰. If we were only assuming that others are conscious, it would be a conclusion. And as such, it would have to be based on evidence and arrived at with argumentation. But what could count as evidence that someone is conscious? A document, a CT scan, a visit to the doctor? Based on what evidence would we be able to draw this conclusion? It is interesting to think about when we would actually ask questions like “Is she (really) conscious?”. If we consider such questions we will see that in general there is no real use for such questions, except in very special circumstances, e.g. when someone has just fainted or is awakening from an operation. Only in such special circumstances it makes sense to ask questions like “Is she conscious or not?” (Cf. PI 416). To Wittgenstein, we are not assuming that others are conscious, but it is a basic fact of our form of live.

Rather than being a belief or hypothesis about other beings, carefully constructed on the basis of experience (and evolution), it is a fact of life that we have learned when we grew up (and fairly early at that). That is, to understand intentional-state concept is to be master of a technique, and part of that technique is that you see other beings as intentional. Seeing them as non-intentional is in fact something we have to put a lot of effort into: on considering imagining that other people are automata, Wittgenstein retorts: “But just try to keep hold of this idea in the midst of your ordinary intercourse with others, in the street, say!”. Clearly, that would be a difficult thing to do.

⁴⁰ Wittgenstein would call this *presupposed* in our lives, actions and activities.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

Thus, we do not act on the *belief* or *hypothesis* that others have a mind: we simply communicate and interact with other conscious beings. It is a basic certainty in our lives that others are conscious like we are, *not* knowledge that has been confirmed by past experimentation. Nothing in our lives would make sense if this certainty was not the same! And this is not a matter of being highly convinced about the truth of the proposition “other human beings have a mind like myself”, it is a fundamental difference between knowledge, where doubt is involved and where there is a need for justification/proof, and certainty, where doubt and justification/proof are useless⁴¹. Wittgenstein writes: “My attitude towards him is an attitude towards a soul. I am not of the *opinion* that he has a soul”. It seems that Dennett, who claims to follow Wittgenstein, has only taken in the first sentence of this quotation, but not the latter.

There is another issue relevant here, namely that it does not make sense to speak of a brain, a thermometer or a computer as if it were conscious entities themselves. That is, the words we use when we speak about conscious entities, such as “belief”, “desire” and “thought” cannot be applied to those entities while meaning the same thing. As Hacker puts it:

One cannot intelligibly ascribe ‘intentionality’ to molecules, cells, parts of the brain, thermostats or computers. Not only is it a subclass of psychological attributes that are the appropriate bearers of intentionality and not animals or things, but, further, only animals, and fairly sophisticated animals at that, and not parts of animals, let alone molecules, thermostats or computers, are the subjects of such attributes. (Bennett & Hacker, 2003, p. 423).

It is only of a normal human being, and several other intelligent living creatures, that we can speak of desires, beliefs or fears without speaking nonsense – or in a non-literal sense. Do we really want to say about a thermometer or a chess-playing computer that it ‘wants’, ‘believes’ or ‘thinks’ and mean the same by these words as we do when we speak of human behaviour? I would say no: to say that a thermostat can think makes as little sense as to say that it can drive a car, go shopping, or calculate. As Wittgenstein writes:

Only of a living human being and what resembles (behaves like) a living human being can one say: it has sensations; it sees; is blind; hears; is deaf; is conscious or unconscious. (PI 281).

⁴¹ This difference is elaborately discussed and explained in Wittgenstein’s *On Certainty*.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

At the same time, this accounts for the fact that in fairy tales or cartoons, inanimate objects have faces, a mouth, speak, are jealous etc. Only when we ‘humanize’ them enough in this way will they begin to meet the special criteria for the ascription of mental states. Wittgenstein continues:

Look at a stone and imagine it having sensations. – One says to oneself: How could one so much as get the idea of ascribing a sensation to a thing? One might as well ascribe it to a number! – And now look at a wiggling fly and at once these difficulties vanish and pain seems to be able to get a foothold here, where before everything was, so to speak, too smooth for it. (PI 284)

The point is that it simply doesn’t make sense to ascribe sensations to inanimate objects because the whole practice of using such words is fundamentally connected with behavioural criteria⁴². Moreover, the fact that we can use the same words to describe the ‘behaviour’ of a thermometer or a computer is not an indication – and certainly not evidence – of these entities being conscious or intentional in any way. Even the word “behaviour” applies in a different way to these entities as it is meant to bring out the physical changes such entities undergo when they are functioning and not the meaningful, social, psychological behaviour of conscious beings. For example, when I say of a rock or thermometer that it behaves such and such I am using a different type of description than when I am telling you about the strange behaviour of my neighbour at the party last night. That is, we are in fact playing a different language game and the techniques in use when describing physical behaviour are different than those in use when describing human behaviour. Only of humans and what behaves like humans we say such things. Both ways of describing have to be learned, although physical descriptions could be said to be a secondary use of social descriptions⁴³.

4.4 Internal versus External relations

4.4.1 Science and Philosophy revisited

This section will focus on the difference between internal and external relations. This difference is bound up with the difference between science and philosophy that we have discussed throughout this thesis a number of times. Hans-Johan Glock, a known

⁴² This does not mean that Wittgenstein is a behaviourist of any sort, but that “fine shades of behaviour” play an important role in the process of learning to ascribe mental states to others, i.e. the process of learning to use the words that we call mental states.

⁴³ For a more elaborate explanation of the idea of secondary sense, see Ter Hark (1990).

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

Wittgenstein commentator, expresses this latter difference as follows: “logic, mathematics and philosophy are concerned with propositions which are *logically* necessary and hence a priori” (Glock, 1996, p. 190) while he describes the empirical sciences as concerned with a posteriori propositions. Friedrich Waismann has been known to discuss issues of logic, science and philosophy with Wittgenstein extensively in the period of the Vienna Circle. In a chapter about the causal interpretation of language, partly aimed against Russell’s *The Analysis of Mind* (1921), Waismann writes the following: “the connection between sentence and fact is a *logical one*, not one drawn from experience” (Waismann, 1965, p. 128). This again shows that no number of experiments can determine the meaning of a word⁴⁴, nor can experiments either prove or disconfirm the internal relation between a word and its meaning, i.e. between representations and what they represent. Rather, the connection between a word and its meaning is a matter of agreement and conventions within the culture it belongs to: it is a built up collection of habits, practices and institutions in which the word performs a certain function. But this does not mean that the meaning of a word cannot be explained to you explicitly, i.e. in terms of definitions. The point here is, however, that the meaning of a word is not something that we establish after a number of experiments but that it is something you have learned as part of your cultural habits, practices, customs, etc. and which has become a certainty to you in daily life. (The difference between reasons and causes can help to explain this as well: you have certain reasons to use a word, but you are not caused to use that word (e.g. by your sensory input). If the latter were the case we *could* conduct experiments and find the ‘true’ meaning of a word, but as said, this is not how representations function.)

Wittgenstein writes: ““The meaning of a word is what is explained by the explanation of the meaning.” I.e. if you want to understand the use of the word “meaning”, look for what are called “explanations of meaning”.” (PI 560). This shows that meaning and explanations of meaning are essentially bound up together. What explanations of meaning are, is illustrated by the following passage, written by Waismann and based on his conversations with Wittgenstein:

The meaning of a word is really given to us in the explanation of this word, and the explanation always consists in specifying the rules for how the word is to be used, perhaps reinforced with examples. [...] What really is a pawn in the game of chess? Is it this particular wooden figure? Or is it the shape of this block of wood? If I am to explain to somebody the significance of the pawn in chess, then I explain to him the rules that hold for this piece. The piece is, as it were,

⁴⁴ That is, if we look at everyday language. In science there have been a number of conceptual changes related to empirical research. Consider e.g. the well-known example of the change of meaning of the scientific concept “water” once the chemical composition was discovered.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

the point of intersection of all the rules that hold for it, and thereby it is defined. [...] The explanation of a word is similar to the explanation of the meaning of a chesspiece. This meaning doesn't lie in the aural form of the word, just as the meaning of the chesspiece doesn't lie in the shape of the wooden block but in the rules. To the question about the meaning of a word a description of the use of the word provides the answer; hence the question was only a question about its use. If I teach someone all the rules that hold for a particular word, I have thereby given him the whole meaning. (Waismann, 2003, p. 453)

This passage also shows that a word, e.g. "red", and its meaning, the colour red, are not independent: what a pawn is only makes sense within the chess game, and given the rules for that and other pieces in the game. What a pawn really is in the game of chess can thus not be known independently of the activity, or practice, of playing chess. The same holds for the word "red": one cannot determine whether something is red independent of understanding, that is, having mastered the use of, the word "red". And this is in turn related to the mastery of other colour concepts, such as "blue" and "lighter" and "darker". In this way, it is thus not experience itself that holds up the meaning of a word, but having learned to use a concept within a number of activities. The meaning of a concept is something that one should relate to conventions and agreements, not to experiences.

In contrast, hypotheses in science can be refuted: it is essential for a hypothesis that it is refutable in principle, if it is impossible to find counterevidence to a thesis, it can not be called a scientific hypothesis at all. Hypotheses can be justified by proof, a chain of arguments based on data that explains them. On the other hand, Wittgenstein's 'grammatical propositions' can be proved nor disconfirmed. They are not refutable in principle; it is not conceivable that they would be refuted. Imagine that we would try to refute the sentence "that is red" when used in the context of pointing to a red rose. Or imagine that we would try to refute that same sentence in a context where there is no red present. Even in that latter case, we could not speak of a refutation but only of a false sentence, or of a wrong use of the word red (depending on the circumstances). The point is that if such sentences were different, i.e. when the meaning of "red" would somehow 'turn out' to be different after careful scientific investigation, our whole system would be different (and not just one part of it). If we change the meaning of the word "red" somehow, this will influence a lot of daily talk – and activities – that concerns colours and even activities in which not the colour but its associations play a role⁴⁵.

Wittgenstein makes an interesting remark on the difference in type of explanation that is offered when he is writing on Frazer's *Golden Bough* – that is, on the difference between causal and rational explanations. Here, he discusses whether we should explain

⁴⁵ Such associations can be called a 'cloud of meanings' and involve things as that red is in our form of life associated with aggression, stopping, danger, warning, evil, etc.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

the impressiveness or frightening nature of certain rituals by means of an historical explanation. He writes the following passage:

But an hypothetical connecting link should in that case do nothing but direct the attention to the similarity, the relatedness, of the facts. As one might illustrate an internal relation of a circle to an ellipse by gradually converting it into a circle; but not in order to assert that a certain ellipse actually, historically, had originated from a circle (evolutionary hypothesis), but only in order to sharpen our eye for a formal connection (RF, p. 69)

Wittgenstein here again brings to light the difference between a causal explanation⁴⁶ and a perspicuous representation when it comes to our understanding of the meaning of something, be it a word, a ritual or a symbol.

Moreover, if we consider the role of certain propositions in human life we get the following argument: propositions such as “this is a hand” or “this is red” can generally function to explain the concept of a hand or red, not to express a piece of (scientific) knowledge⁴⁷. We have not discovered our hands nor have we discovered the colour red when we learned these expressions, rather, we have learned to use these concepts. By having learned them as part of human activities, i.e. as part of our language games, they have in turn constituted what ‘hand’ and ‘red’ are. The relation between the activities in which the word “hand” plays a role and what the word “hand” means is internal, not external. This means that a hand and the word “hand” are not independent: it is not possible to distinguish something as a hand without having mastered the use of the concept “hand” (or a similar concept in a different language) and vice versa. Rather than having arrived at such propositions after investigations, our investigations *start* from such propositions. If we hadn’t already mastered numerous language games, techniques and practices we would not have been able to ask empirical scientific questions at all. The meaning of the word “hand” is thus determined by practices, institutions, rules, habits, ceremonies, etc. in our form of life, in which the word “hand” plays a role. The focus here on behaviour and practice might give the impression that the meaning of a word is something empirical after all, but the point here is exactly the opposite. The relation

⁴⁶ The historical explanation here counts as a causal explanation since it tries to specify which factors had which effects, i.e. it gives an explanation in terms of (lawful) cause-effect relations.

⁴⁷ Consider what an error would be with such expressions: when would we say that the sentence “this is a hand” is an error? When scientific investigation discovers that what we were talking about is in fact an intricate replica? Probably yes, but then again, do such events occur in general? And does that amount to the type of error we are looking for here? I would say not. Normally, an “error” here would be an incorrect use of the expression, not a fake hand (that would be deception: we could not actually hold a person responsible for believing it is a hand).

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

between the word “hand” and all the activities it is active in is internal, since neither the activities nor the concept itself can be investigated independently. That is, in order to know what “hand” means one needs to master the various activities, and in order to understand the activities themselves one needs to master the concept “hand”. They go, so to say, hand-in-hand. It would thus misrepresent Wittgenstein’s thought here to say that understanding a concept is in the end a matter of experience, and that the criteria for understanding is purely behavioural (cf. McGinn, 1997, p. 93).

An example might help to show how representations can obtain their function within a community using that representation within their practices and daily activities. Consider the following symbols:



With some consideration, it is obvious that these symbols do not represent intrinsically, that is, there is nothing in these symbols that make them represent “stop”, “play”, “pause”, etc. If someone were to say that the fact that the play-symbol is aimed forward as we read from left to right, would that count as an explanation? It *does* have a relation with our culture, but the mere fact that in other cultures people read in different directions shows that it is not an intrinsic feature of the symbol itself, but rather a cultural convention. The word “convention” is slightly misleading here, as it might give the impression that we are talking about an explicit agreement. This is not the case, it is only meant to bring out the fact that it is a contingent aspect of the symbol’s meaning and not an intrinsic feature. Moreover, such an explanation of why the play-symbol means “play” is exactly the type of historical/causal explanation Wittgenstein tries to get away from in the above quoted passage.

Rather, what we should realise here is that these symbols have (only recently) been introduced to our culture and we have all adopted them for a special purpose. The fact that the play button *means* “play” to us, is constituted by all the different activities in which we (have learned to) use this symbol. Its meaning is thus very clearly determined by how it functions in our society: if you press a button with the play-symbol, a movie will start, a song will play, or something of that kind. This symbol is not a ‘mere’ triangle anymore, but it has become an entirely different symbol⁴⁸. Note that even a ‘mere’ triangle is conceptually loaded. This point just aims to bring out the fact that the play-button does not *mean* “triangle” but something different⁴⁹.

⁴⁸ Note that this is what Dretske would call a RS of type I: it derives its representational powers from our way of using it. See Section 2.4.3.

⁴⁹ NB: I will elaborate on the aspects of learning, training and the community in Section 4.5.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

4.4.2 Problems with the indication approach

Dennett and Dretske both speak of brain-states – i.e. the physical states or structures and patterns in the brain – as having a meaning or as exhibiting aboutness. Thus, representations do not have to be exclusive to the brain to them, as any physical state can be candidate for representational powers. Subsequently, a large number of entities can be called representations according to Dennett and Dretske⁵⁰: both speak of the ‘internal state’ of simple systems such as thermometers, galvanometers, opnioid receptors or chess computers as being a *representation* of its surroundings. Such internal states represent the environment by means of the causal and nomic relation between the thing they ‘monitor’ and the internal state of the system, e.g. between room temperature and the internal condition of the thermometer⁵¹.

Presenting reference or meaning as indication results in a number of problems. First of all there is a problem with misrepresentation or error (cf, Dretske, 1988, pp. 64-70). How can, for example, a thermometer misrepresent the room temperature? Dretske replies that it can’t: if it is hardwired wrongly, that is not really what error amounts to, as he himself notes (Dretske, 1988, p. 65). A thermometer is a relatively simple intentional system and things such as misrepresentation (e.g. thinking about a unicorn or that it is raining when it is not) are quite complicated matters that come in later, i.e. that come in only at what Dretske calls Representational Systems of type III. Only these complex and highly evolved systems of representation can have the sort of representations that are able to misrepresent (thus, for example, knowledge is something that only the higher animals can have for Dretske).

While this is an interesting subject in itself, I will have to leave this issue to the reader, as there is a more fundamental problem with the indication approach: presenting meaning as indication is misconceived because it presents meaning as an external relation instead of an internal relation. In the previous section we already saw that what a word means is bound up with how we have learned to use it. What “red” means is bound up with how our community of language users employs that concept, and how they teach it to their youngsters. The meaning of this concept is bound up with a language user’s

⁵⁰ It has to be mentioned that Dretske is more subtle than Dennett on this account, see Section 2.4.3

⁵¹ Dretske is far clearer on this point than Dennett. In fact, Dennett seems to evade this point on a number of occasions, but from some of his remarks we can reconstruct a position such as the one Dretske is constructing. It is hard to be sure whether Dennett actually has an indication approach like Dretske, but he does call representations ‘mirrors’ of external conditions and sees them as conditions that control output behaviour because of that mirroring faculty (see Chapter 2). However, even if Dennett disagrees with the nomic nature of the relationship between a representation and what it represents, he is still trying to find a causal relation between the level of psychological descriptions and physical descriptions, which is at odds with what Wittgenstein is aiming for (a conceptual rather than an empirical/ontological relation).

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

participation in a particular form of life. It is thus rather a matter of convention and agreement, and not a matter of build up experiences. A concept gets its sense within the complex form of life: the way language users live, what they do, the way they respond, the activities they employ and so on. The same holds for mental state concepts such as “pain” or “desire”. This means that not all concepts represent in the same way. The danger here lies in thinking that all concepts, including “thought” or “desire” are bound up in a particular way with what they represent, i.e. a way that for example describes or pictures what is going on – in this particular case: in the brain. Take the following example: stating that you need a glass of water is not a description of the dehydrated state of your body, but it is a way of expressing that you are thirsty. To Dretske, a sentence such as “I believe I need a glass of water” can be seen as a description of the dehydrated state of your body. And for Dennett, the expression of such a sentence can be used to predict the behaviour of the system expressing it: if that is successful, the system can be called a “true believer” (Cf. Dennett, 1987, p. 29)⁵². But to Wittgenstein, it is not a descriptive or predictive sentence at all: it is how we express the fact that we are in such a condition, but it is not the same as the physical condition of your body nor is it a description of it. You have learned to employ the relevant words in such a manner that you can e.g. express your thirstiness or obtain a glass of water. In this way, sentences such as “I believe I need a glass of water” and “I’m thirsty” become part of the criteria of someone being thirsty. Such sentences in turn determine what it is to be thirsty, i.e. they determine the concept of thirsty. The relation between “thirst” and the meaning of that word (i.e. what you can achieve with it) is an internal relation: what concept X means is dependent on how you have learned to employ X in daily life, i.e. on what you have learned to do with that concept. Thought and referent are not two independent ‘things’ that we – as it were looking from a position outside and independent of language and the world⁵³ – can relate to each other, but it is an internal relation. What it means to think

⁵² Note that to Dennett, almost all sentences of this type could be preceded by “I believe that”: If I were to say that I’m hungry, this is in fact the same as stating that I believe that I’m hungry, and so on. To Dennett, such a sentence thus has predictive, and therefore explanatory, powers. To Wittgenstein such a sentence is a performative act: you express a certain attitude. From this expression someone else could indeed conclude that you would like food (or water), but this is not a ‘prediction’ in the scientific sense of the word. Rather, it is the consequence that is associated with the expression of such a sentence. Note also that in contrast to Dennett, Wittgenstein observes an asymmetry between first and third person utterances in the *Philosophical Investigations*, e.g. concerning the concept “to know” (cf. PI 249 and further).

⁵³ This is problematic because we cannot view the world independently from the way we have learned to deal with it, a way that is bound up essentially with our mastery of language. To repeat something discussed earlier: what *red* is, is dependent on how we employ the word “red” in all our practices, which in turn determines how we teach that concept to our youngsters. So, our view of the world is

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

about X is not that you have that object ‘in mind’ but it has to do with how you have learned to use X as a concept. The ‘content’ of X is inseparable from the word itself although it is not something that comes intrinsically with it.

An important difference that can aid in seeing the difference between an internal and external relation from another angle is that between a sign and a symptom⁵⁴. This difference will demonstrate that representation is not the same as reliable or lawful indication. In the following illuminating passage⁵⁵ Waismann introduces a number of examples:

A further confusion comes from the use of the expression “sign”, “meaning” and everything of this sort. For example, one says, “Red is a sign for this → colour”, “The low barometer reading is a sign that it will rain”, “The swallows are flying low, this announces rain”, “This cloud means rain”. The words “a is a sign of ...”, as we understand them, are always an explanation of the sign a. By contrast, if one calls the barometer reading a sign of rain, this then is a communication about a causal connection between the barometer reading and rain. That is, the words: “The sign a means this → object” or “a = b Def”, which are explanations of signs, say nothing about causal connection between two events. That is, it is nonsense to say this explanation would be confirmed or disproved by future experience. The proposition speaking of the barometer reading is a hypothesis; the explanation “a is a sign of ...” is an agreement in our sense. Of a hypothesis we can say that it will be confirmed or disproved by experience, but not of an agreement. An agreement can be broken, or one can go by it. The explanation of a sign is always arbitrary. (Waismann, 2003, p. 89)⁵⁶

not independent of our mastery of language, and it is thus impossible to inhibit a position that *is* independent of language (and the world).

⁵⁴ This difference is similar to Grice’s distinction between natural and non-natural meaning, a distinction that causalists often wish to refrain from (Grice, 1957).

⁵⁵ From Notebook I, published in *The Voices of Wittgenstein: The Vienna Circle*, Waismann, 2003

⁵⁶ Cf. the following passage: “The fluctuation in grammar between criteria and symptoms makes it look as if there were nothing at all but symptoms. We say, for example: “Experience teaches that there is rain when the barometer falls, but it also teaches that there is rain when we have certain sensations of wet and cold, or such-and-such visual impressions.” In defence of this one says that these sense-impressions can deceive us. But here one fails to reflect that the fact that the false appearance is precisely one of rain is founded on a definition.” (PI 354). Here Wittgenstein shows that “a low barometer reading *means* rain” denotes a causal connection (like your sense-impressions as a result of being in contact with rain) and not representation.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

Let us investigate this. Even though we call clouds or a low barometer reading a “sign” of coming rain, this is not the same kind of “sign” as a word, signpost or diagram. Clouds, the low stance of a barometer, the ‘internal state’ of a thermometer, etc. are not *signs* in that sense, they are *symptoms* of the physical situation that they are part of. But to Dretske, this relation of indication is the basis for representation. A relation of indication itself cannot be wrong according to Dretske (cf. Dretske, 1988, pp. 64-70), which is in line with its causal nature, but somewhere the transition between purely indicative systems and truly representational systems is made. Note also that Dretske nonetheless insists on calling e.g. fuel gauges or thermometers “representations”. For example: “My fuel gauge is not only a representation *of* and empty gasoline tank; it is also (when things are working right) and empty-tank representation. That the tank is empty is what it indicates, the information it carries, the comment it makes, about that topic.”⁵⁷ (Dretske, 1988, p. 71). Again we see that the surface grammar of our language can mislead us. Recall that the causal theory of representation uses examples like these to make the argument for the relation of meaning being causal in nature. But like the barometer, a thermometer indicates a certain temperature because it is *part of* the room and *part of* the physical and causal process of heating (temperature). It indicates something about the physical situation it is in, because it *is* that physical situation. Of course, the fact that we use barometers and thermometers to read off such ‘information’ about our environment is a cultural phenomenon in itself that needs to be learned. In fact, what concepts such as “temperature”, “heat”, “warmer”, etc. mean is in part determined by practices such as measuring the temperature with a thermometer. But this relation of (causal) indication is *not* the same as representation.

At this point we can see how the difference between a sign and a symptom relates to the difference between internal and external relations. We have already seen that the type of relation that holds between a thermometer and the room temperature, a causal relation, is an external relation. We now see that this type of relation involves symptoms: just as having red spots on your skin is a symptom of having the measles, or just as the low milk-production of sheep is a symptom of bad feeding patterns, likewise the rising or falling of the level of quicksilver in a thermometer is a symptom of the room temperature. And we have learned to use that symptom within our community to do a number of things. In contrast to this, we see that the ‘relation’ between a word and its meaning, e.g. the relation between “red” and *red*, is of a different nature. Whereas a thermometer and the room temperature are not conceptually dependent on each other, the practice of using colour words and “red” are. The latter are internally related whereas the former are not.

⁵⁷ The difference alluded to in the quotation is not relevant for the current discussion, but the quotation shows Dretske’s way of talking about a representation (sign) in the case of a symptom.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

To sum up: a symptom is externally related to what it is a symptom of, e.g. a barometer will indicate a certain weather type because the pressure is low or high, which is directly related to that weather type. A sign on the other hand, is internally related to what it is a sign of, for example the word “rain” is dependent on how we (have learned to) use that word, i.e. in which of situations it is applicable. The word “rain” is thus not independent on that which we call rain. Mixing these two relations up, as I think both Dennett and Dretske are doing, results in a wrong view of how representation works.

4.5 A Positive Outlook

The final part of this chapter will concentrate more on the positive aspect of Wittgenstein’s philosophy. It has to be noted that the positive and negative characterization of Wittgenstein’s philosophy given in Chapter 3 is not as strict as it might seem. In fact, most of Wittgenstein’s positive points, i.e. points that show how one should think about or look at representation, are intricately bound up with his criticism. For example, criticism of mixing up signs and symptoms, or internal and external relations, goes hand in hand with explaining these distinctions. Thus, when Wittgenstein criticizes someone for not making this distinction properly, step by step he will explain the distinction as well. This means that the positive outlook is there in the text, but needs to be built up partly from the criticism Wittgenstein has developed, while at the other hand it can be reconstructed from remarks spread out through his work. Some aspects that can be categorized under the positive aspect of Wittgenstein’s philosophy have thus already been discussed at this point. In the remainder I shall concentrate on the points that deserve extra consideration or explanation.

4.5.1 What it is to understand a concept

As stated before, Wittgenstein is battling the urge to look for the *essence* of a concept as the explanation of it. When one understands what a concept means, it does not follow that one has discovered a pattern or that one reasons from a set of observations towards a concept or interpretation. It is in fact the urge to look for the essence – i.e. the theoretical attitude – that we must rid ourselves of in order to understand the true nature of our problems and the solutions to them.

An example will help us to clarify this. Consider the following: we all know what music is and which things can’t be called music. We were not born with this knowledge, but we learned it when we grew up. For example, when we hear random noises we would not classify this as music; when someone drums a rhythm on the table with her fingers, we can call that music; when someone whistles a tune, this can be called music; when we

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

hear a car driving by, we would not call that music. Basically, we all know quite well which things count as music and which ones not. That is, we all know quite well how to apply the concept “music”; we know when the concept is being used wrongly and we can also explain the concept to someone else. Note that even without being able to – or feel the need to – specify particular explicit constraints for what counts as music or not, we are already able to judge whether something is music or not. Moreover, even if we *would* find a number of sufficient and necessary constraints that can set out music apart from all other types of sound, those constraints would not be the constraints we actually use as the rules for applying the word music, i.e. it is not this set of constraints that make up the concept of music. As the case is, it is probably very hard to specify a number of constraints that can in reality distinguish music from non-music: some types of music are hardly rhythmical whilst others are very constant; some types of music use no regular instruments at all, whilst others only use drums and guitars, and so on⁵⁸. The diversity is immense. If we could find such constraints (maybe by using AI applications such as neural networks to mine the data for us) these could be used by computers or robots to set music apart from other sounds, but the point is that we don’t use those features to ‘detect’ whether something is music or not.

Rather, the concept is determined by our common ability to judge whether something is music or not. And this ability is something that is common to our way of life, something that we have learned to employ within the same cultural, social and historical contexts. It is because of the fact that we share the same practices and methods in which we employ the word music (e.g. playing instruments, talking about bands, going to concerts, etc.) that this word means the same to each of us.

Moreover, defining the concept – making a definition of it either by description or by trying to figure out its constraints – is something that is only possible after having mastered the application of the concept itself. Only if you already know what a concept means, you can give a viable definition of it (which, again, will not be a sufficient and necessary specification of it). Surely, once we have mastered a large number of concepts and the technique of introducing and understanding new concepts by giving definitions of it, we will be able to *learn* new words from definitions (e.g. from the dictionary) – even though this is different from giving definitions of concepts ourselves. But foremost, this is also a technique we must master. This example in addition shows that with Wittgenstein’s method we are not merely talking about the word “music”, but we are in effect investigating what music itself is: we come to see what kind of answer is suitable

⁵⁸ This shows that “music” is a family-resemblance concept: if you compare two examples of music you can probably find something these have in common, but if you take in another one, chances are that this one will not have that same feature in common. What we call music is combined like a thread: all the fibers are connected but there is not one fiber running through the whole thread (cf. PI 67).

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

for the question “Why do we call certain things music and other things not?”. The answer is not some essential feature or pattern, but lies in coming to understand our use of the concept in everyday life. The answer is already there: one has only to look at all the (correct and incorrect) applications of the word “music” to see what it is about.

The example thus shows that recognizing music, i.e. understanding the concept music, is not the application of a series of sufficient and necessary constraints that delineate the border between music and non-music. Rather, understanding “music” is the result of having learned a number of language games in which we employ the word. Indeed, to be able to play the language games we need a sophisticated system that is able to recognise patterns, as Dennett says, but this pattern-recognition is not the basis for our use of the concept. We don’t call something music because it resembles a certain objective pattern that is to be found in *all* music but we call something music because we have learned to use the word in a specific practice for specific purposes. In other words: we have good reasons for using the word “music” in the situations rather than being caused to use the word by a certain external/objective pattern our brain picks up⁵⁹.

We could also say that a word is surrounded by an ‘atmosphere’, and one who is familiar with this atmosphere understands the application of the concept (cf. PI 117; p. 181). This atmosphere is however not something that goes along with the word no matter what, but it is rather connected with the environment in which the word is used. So the atmosphere of the word “music” consists of all the situations in which that word would be applicable. This atmosphere is what we have to get acquainted with in order to understand the word and all the subtle differences between it and similar words: then we will understand that e.g. “angry” is different from “enraged” or “furious” and which word is the most suitable for which type of context. This also explains why we can’t simply take a word from one language game and start using it in another: the word and its employment in its context are bound up together in an inseparable way.

4.5.2 Learning a language is learning a form of life

One of the most important points we can take from Wittgenstein, is that learning a language is not something like learning labels for objects, experiences or patterns that you already are familiar with, but rather that learning a language comes down to learning a point of view – a *weltanschauung*. It is important to notice that what a word means is hereby a matter of the community: it is only within a community that a consistent technique of using words/signs within practices can thrive. This doesn’t mean that one

⁵⁹ The difference between such a pattern and what e.g. Dennett is looking for (Dennett, 1991b), is that we are not speaking about a spatio-temporal pattern of features in the music itself, but rather about a pattern that holds within the language use of the community, a wide and flexible pattern that consists of agreement and judgements rather than specific sufficient and necessary conditions.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

person in physical isolation cannot use language or other signs, but it means that one has to grow up within a community and learn the practices there. Obviously, an (imaginary) example like Robinson Crusoe shows that one can perfectly use language correct and meaningful in physical isolation. The point here is that language and all the practices and techniques it is weaved together with is an immense cultural phenomenon that is held up by an entire community. This is what Wittgenstein calls a *form of life*.

But at the same time, it should be clear that there are large differences between such communities, be they entire countries or merely a group of peers such as carpenters. The way words and other signs function within such groups can be very different. In considering this the fact that we use the word “red” to distinguish red things from differently coloured things seems to be a somewhat arbitrary matter. There is nothing in the word “red” itself that makes it represent red rather than any other colour, or for that matter, rather than anything different at all (e.g. we could also imagine that “red” is used in the same way as we use “hello” or “stand over there”). However, that we use “red” to denote red is not an arbitrary matter: it is a contingent matter. Even though it could have been different – and even though we can change it if we want to – it is at this point important which particular sounds/letters you use. One could say that the start or beginning of the use of such a word is arbitrary, but not that the current use of the word “red” is. Because we have all learned to use this word in the sense it has, within our form of life, has become a matter of fact for us and is thereby a grammatical rule for the way we use language.

When I write that the ‘start’ or ‘beginning’ could have been arbitrary, I meant the following: the meaning of a word is dependent on the role that we assign the word in our language games. That is, we could teach a child a different use of the word “red”, if we were consistent enough. What a word means is determined by all the language games in which it is used meaningfully. For example, to use “red” to distinguish red objects from blue objects, to explain to someone the use of the word red by showing red objects and saying “this object is red”, to stop in front of a red stoplight, to understand what it means if someone says “his face turned red”, etc. There are many different language games that use the word red, but they are not all connected by a (set of) common feature(s). Rather, they are related in the way of a family resemblance⁶⁰. It is because we have all mastered these different language games – in which “red” has a function – we understand the word “red”. The meaning of this word is thus not made up by a single relation between e.g. a set of sounds or marks on paper on the one hand and a frequency of light on the other, it is made up out of all the different functions that we have given the word in our language games. To speak of a relation here can already be quite misleading, because relations generally obtain between two independent items, i.e. the relata can be identified

⁶⁰ See Chapter 3 on methods for an explanation on the notion of a family resemblance.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

independent of each other⁶¹. But the word “red” and red itself aren’t independent in that manner: what red *is*, is determined by your mastery of the use of the word “red”. It is human practice, and not a flow of information or lawful indication, that endows words and thoughts with meaning. Or, in different words, it is something internal to language and not external to it that constitutes its meaningfulness.

An interesting comparison at this point is to consider ceremonies in human culture. In *Remarks on Frazer’s Golden Bough* Wittgenstein discusses a number of such ceremonies and, shows how they get their meaning from the role we assign to certain acts rather than from an intrinsic feature that such acts have. This is clearly visible in the following example:

Recall that after Schubert’s death his brother cut some of Schubert’s scores into small pieces and gave such pieces, consisting of a few bars, to his favourite pupils. This act, as a sign of piety, is *just as* understandable to us as the different one of keeping the scores untouched, accessible to no one. And if Schubert’s brother had burned the scores, that too would be understandable as a sign of piety (RF, p.66).

Here we see that both acts could be conceived of as good and in memory of Schubert. Both acts can be made sense of in a positive manner, as long as we surround the act with the correct background story. Here we see, that the reasons that are provided for the acts are decisive, not the acts themselves. If we tear up the paper and distribute them amongst people, this could be because everyone will in this way have something to remember Schubert by⁶². Likewise, if we would burn the notes that could be a symbolic act as well, which can leave a lasting impression on the bystanders.

Consider also the case of the King of the Wood of Nemi. Frazer tries to explain why this story is so dreadful by looking at the history of the ceremony. Hereby, he tries to establish an external connection between the story, and e.g. the horrible nature of killing. But as Wittgenstein points out when he considers a similar case – the Beltane Fire Festivals – the dreadfulness or sinister does not stem from the *origin* of the practice (RF, p.75). For what the particular origin is, remains a hypothesis that could be otherwise, and

⁶¹ This difference can also be explained as a conceptual versus an empirical relation between entities. Two colliding billiard balls are empirically related, but “bachelor” is conceptually related to “unmarried man”. This difference is also reflected in the well-known division between analytic and synthetic truths in Empiricism.

⁶² It has to be noted that this is only one of several meaningful rational explanations that could be given of these acts. This example simply serves to show that no matter what the act is, we can come up with a story that can turn the act into something good or evil depending on the story, i.e. the reasons we provide.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

we do not want to say that the ceremony is so dreadful only if its origin is such-and-such. As Wittgenstein says: “I want to say: The deep, the sinister, do not depend on the history of the practice having been like this, for perhaps it was not like this at all; nor on the fact that it was perhaps or probably like this, but rather on that which gives me grounds for assuming this” (RF, p.77).

And what gives us the grounds for assuming this is the fact that we already treat such stories as dreadful. The story itself isn’t dreadful, we could conceive of a way of dealing with the story that makes it into a mere historical summation of fact or maybe even a happy story, all depending on the context/atmosphere we surround it with. We should concentrate on the fact that we *do* regard the story of the King of Nemi as dreadful. As Wittgenstein points out, Frazer already tells the story in a dreadful tone of voice. He takes the story to be a dreadful one. Not because dying in itself is something dreadful, but because we treat stories about death as dreadful (in general). The dreadfulness comes from our way of dealing with it: “the deep and the sinister do not become apparent merely by our coming to know the history of the external action, rather it is *we* who ascribe them from an experience of our own” (RF, p.77). Wittgenstein emphasizes that it is our way of dealing with things that gives a set of actions their ceremonial value.

In a way, it might now seem that “every view has its charm” (RF, p.71). But Wittgenstein warns us not to take this statement as a relativistic doctrine: what he means is that magical and religious practices, or ceremonies, are not the type of thing that we can call either right or wrong. Wittgenstein is not defending that we can talk everything right or wrong. What he means is that it simply doesn’t make sense to call such views wrong⁶³, whereas it does make sense to call them significant for some people. In some cultures or language games certain acts (or words) have a significant role. Just like we used to bury people when they die, Indians burned the bodies of their dead. Just like some cultures tend to mourn in black when a loved one dies, other cultures celebrate their entrance to the afterlife. Rituals of mourning or celebrating the dead are not good or bad intrinsically, but get their significance from the way we deal with these acts.

The analogy here is that we can say the same things about words: they do not have their meaning attached to them, it is not intrinsic to e.g. the word “red” that it denotes red. Wittgenstein attacks this idea in PI 117, where he states it as follows: “As if the sense were an atmosphere accompanying the word, which it carried with it into every kind of application”. He replies that in order to understand why a sentence makes sense to us, we should ask ourselves “in what special circumstances this sentence is actually used” (PI

⁶³ Cf. PI, p.227: “Is a coronation wrong? To beings different from ourselves it might look extremely odd”. This shows that a coronation is a custom and that it derives its meaning from how the community uses it – not from the intrinsic properties of the activity itself.

CHAPTER 4 – A CRITIQUE OF THE CAUSAL THEORY OF REPRESENTATION

117)⁶⁴. This is where the word derives its sense from. Note that sense is not equated with use; the meaning of a word is not its use. Rather, Wittgenstein wants to bring out the fact that it is only *in* use that words get their meaning: “Every sign *by itself* seems dead. *What* gives it life? – In use it is alive. Is life breathed into it there? – Or is the *use* its life?” (PI 432)⁶⁵. This remark is supposed to bring out the fact that when we consider language by itself, as it were outside of their position in our lives and activities – i.e. outside the stream of life – a word has become empty and meaningless. This is why Wittgenstein writes “philosophical problems arise when language *goes on holiday*” (PI 38). It is only when we start to consider how language in itself, i.e. sounds and marks on paper, relates to the world that our philosophical puzzle of representation comes into existence. Rather, words are like tools (PI 11) that we employ in our daily practice. It is because of the fact that we use the word “red” in the specific ways we do, that it comes to denote red for us. The meaning of the word “red” is fundamentally related to our form of life.

⁶⁴ Cf. PI 595: “It is natural for us to say a sentence in such-and-such surroundings, and unnatural to say it in isolation. Are we to say that there is a particular feeling accompanying the utterance of every sentence when we say it naturally?” Wittgenstein is obviously showing here that there is no such feeling and that the feelings accompanying your utterances are not their meanings. See also PI 607 where Wittgenstein discusses a similar case of ‘particular atmosphere’.

⁶⁵ Cf. the discussion on language games and form of life in chapter 2

“The riddle does not exist. If a question can be put at all, then it can also be answered.”

- Ludwig Wittgenstein

Chapter 5

Conclusions

5.1 Introduction

In this final chapter, I will provide a summarization of the main discussion in this thesis and focus on its results and discuss the significance of the previous chapters with respect to the research question posed in Chapter 1. I will also give a number of recommendations for future research, as a number of new issues and questions came to mind during the writing of this thesis, but not all of these could be included in the end. Therefore, a number of suggestions related to the issues discussed here will be outlined in the second part of this chapter.

5.2 Research outcome

The research question was:

- (1) “What is the nature of representation?”
- (2) “What critique can be given on contemporary philosophical theories on representation using Wittgenstein’s remarks?”

Let us take a look at what we have undertaken in this thesis to answer these questions. The first question is called the problem of representation, although it has many other forms as well, e.g. “how does a word manage to represent its meaning” or “what is the nature of the relation between a thought and what the thought is about”. This problem is explained in detail in Chapter 1. In the same chapter, I provided an overview of the views of Dennett and Dretske in the same chapter. This overview showed that both philosophers belong to the same tradition. This tradition adheres to what I have called the causal world

CHAPTER 5 – CONCLUSIONS

picture: the idea that the world and all its phenomena can be described in terms of causes and their effects. Despite differences in the views of Dennett and Dretske, their writings have been shown to use this causal world picture as the background against which they have developed their views: both Dennett and Dretske operate within the framework of the causal world picture.

Dennett and Dretske have formulated the problem of representation in the same way: “How can a mere mechanism, such as the brain, be capable of representations?”. They also extend the meaning of the concept of “representation” in the same fashion, encompassing a range much wider than our everyday language admits. Amongst others, we can find them claiming that thermostats, amoebas, brain-cells etc. have internal states that are representations of the outside world and that have the function of steering the behaviour or the system. They call such systems ‘intentional systems’. Even though they differ on their definition of how we can determine what an intentional system is, the point is that they ascribe intentionality or aboutness to these systems.

In Chapter 3, I then set out to give an overview of Wittgenstein’s method of philosophy. This method is fundamentally bound up with Wittgenstein’s unusual style of approaching problems. I have shown that the rationale for this method is in part due to Wittgenstein’s views on the difference between science and philosophy⁶⁶. The difference between science and philosophy comes down to a number of points: there are no theories in philosophy; philosophy does not contain hypotheses; philosophy explains nothing (in the manner that science does), it only describes what is already there; philosophical problems are not empirical and thus experiments will not help to solve them; philosophical problems can only be solved by getting a clear overview of the way language functions (a perspicuous representation). These points are discussed at length in Chapter 3 and return in Chapter 4.

Chapter 3 was also concerned with two characterisations of Wittgenstein’s philosophy: a negative and a positive one. Both roles are connected to one another, as the critique Wittgenstein offers (negative role) often consists partially of showing alternatives and differences, posing questions that will attend the reader to unnoticed aspects, eventually providing what he considers contributal to the solution of philosophical problems, i.e. a ‘perspicuous representation’ of concepts (positive role). One could say that whereas science provides a ‘quick’ and systematic answer to a problem, philosophy,

⁶⁶ Although I have called both Dretske and Dennett philosophers throughout this thesis, one could argue that against the background of Wittgenstein’s views on what philosophy is, they could for example also be categorized as psychologists, anthropologists or linguists. However, what is currently widespread regarded and practised as philosophy does not meet the standards of what Wittgenstein delineated as philosophy. Therefore, I decided to continue to call them philosophers, even though I subscribe to Wittgenstein’s fundamental distinction between science and philosophy. This has been discussed in Chapter 3 as well.

CHAPTER 5 – CONCLUSIONS

in Wittgenstein's eyes, should aim at slowly ridding someone of his philosophical problems, by a type of conceptual treatment. This has been called the 'therapeutic approach'. An important aspect to notice is that Wittgenstein believes most philosophical problems to be caused by an essentialistic craving for theories, hypotheses and idealisations, a craving that is not in place within the domain of philosophy (but can be very fruitful outside of it, e.g. in the natural sciences). Moreover, philosophical problems can be traced back to a misunderstanding of our language on account of this theoretical attitude: if we generalise from a number of examples we are at risk of creating an essentialistic solution where none is to be found: for example, if we compare sentences such as "Peter is angry", "The Morningstar is the Eveningstar" and "That rose is red", we can see that not all uses of the words "is" can be explained by one common factor. The same holds for sentences such as "I hit him because my arm-muscles contracted" and "I hit him because I was angry with him": the word "because" in our language cannot be explained by a single paradigm, such as causality.

The two distinct philosophical approaches to the problem of representation discussed in Chapter 2 and 3 provided the theoretical background for the discussion of our main subject, representation. In Chapter 4 the two approaches come together. The initial part of this chapter was aimed at bringing out the differences in methodology between Dennett and Dretske on the one hand, and Wittgenstein on the other. As noted above, to Wittgenstein philosophical problems are fundamentally different from empirical/scientific problems. This is not the case for Dennett and Dretske. In part, we could sum up Wittgenstein's critique by saying that philosophers alluding to the causal theory of representation have succumbed to the problematic theoretical attitude, and are being misled by the forms of our language. The two main misunderstandings I attributed to the causal theory of representation are: (1) thinking that reasons and causes are somehow of the same species, or that one is a subspecies of the other; and (2) thinking that representation is a relation that holds between two logically (or conceptually) independent items. These two problems form the main answer to the second, specific research question stated above. I have argued in detail why these two points are mistaken aided by Wittgenstein's remarks.

In both cases we can say that the problem is caused by a misunderstanding of the forms of our language, i.e. Dennett and Dretske are misled by the surface grammar of certain expressions. These misunderstandings affect their methodology in investigating the problem of representation. First off, taking rational explanations for causal explanations results in an erroneous view of what mental states are. This in turn lead towards a mistaken view of human behaviour and its interpretation, such as Dennett's intentional stance. Secondly, trying to specify an external relation where in fact an internal relation is needed, resulted in the wrong type of relation specified between the relata (e.g. a word and an object, or a thought and what it is about) Moreover, this latter

CHAPTER 5 – CONCLUSIONS

misunderstanding will also result in a wrong idea of how language and the world are related, i.e. of how representation functions, such as Dretske's indication approach to representation. The solution for both problems was to investigate and clarify these concepts, i.e. to obtain a perspicuous representation of them. It is by investigating different examples and pointing out differences that Wittgenstein leads us towards such a perspicuous representation.

Finally, after having discussed these problems and after having given my view of Wittgenstein's critique to them I turned to a more positive outlook. Even though Wittgenstein's critique partly consists of showing alternatives, examples, exceptions and differences, and already provides the reader with a number of enlightening questions and thought-experiments, there remained a number of things to be said. Instead of trying to find the *essence* of a phenomenon or concept, Wittgenstein shows us the alternative route: that of creating a perspicuous overview of the concept. When one has obtained such an overview, one is capable of seeing the role or a concept within the totality of frameworks it is functioning in. The meaning of a word and the way it is learned is not constituted by certain mechanisms, but is related to the form of life it is active in. A concepts meaning is fundamentally bound up with the way we employ it in our lives, it is the totality of functions the word has in the different language games. That is, a word and its meaning are not logically independent of one another.

Finally, let us summarize this into an answer to the general research question posed at the beginning of this thesis: "What is the nature of representation?". It has to be noted that Wittgenstein's method and views do not lend themselves for an 'easy' answer to this question, as this thesis has hopefully been able to convey. But, to play the devil's advocate for a bit, I will still give a short answer, and hope that the reader will be able to understand that against the background of all that has been said: the representation of a concept is constituted by all the activities or practices it is employed in within a given culture or society. The meaning of a concept is related to many other concepts and is apt to change over time. Moreover, even the concept of representation itself cannot be pinpointed to a single paradigm as especially causalists are inclined to assume. It is essentially a family resemblance concept, meaning that not all the things we call representations share one and the same property. The meaning of a concept is, on account of its being used, the totality of its functionality in a given form of life. As Wittgenstein said: "the meaning of a word is its use in the language" (PI 43)⁶⁷. This means that the culture, and its activities, rituals, habits, needs, crafts, institutions, etc. and the concept

⁶⁷ Note that this quote has been misread or misused many times, e.g. when philosophers try to equate the meaning of a concept with its literal pattern of use. At this point it should be clear that that is not the correct reading of this remark. Moreover, Wittgenstein precedes this remark with the following line: "For a *large* class of cases – though not for all – in which we employ the word "meaning" it can be defined thus:".

itself are fundamentally bound up together: the one cannot be considered independent of the other.

5.3 Suggestions for future research

One of the main arguments in this thesis has been that reasons and causes are not of the same kind. What would be interesting to develop further, is what we should make of cases or situations where reasons and causes are both used at the same time. If we analyse such cases, will we find that sometimes the reasons are actually causes or vice versa? Will each situation require its own approach or can we give a more general account? An interesting example are the theories of Freud, who has combined talk of reasons (motives) and causes into one framework that seems to work as a (soothing) explanation for many people. What is the attractiveness of such explanations? Amongst others, Jacques Bouveresse has written substantially on that (Bouveresse, 1995) and this would be an interesting direction to pursue.

Another point where the difference between reasons and causes comes into play is the reductionist programme that many contemporary analytical philosophers seem to have joined. That is, to what extent can the critique developed within this thesis be applied and extended to the larger framework of the reductionist programme? Obviously, this programme has many connections with what I have called the causal world picture. Moreover, the causal theory of representation can also be placed within the reductionist programme as it aims to find a way to reduce talk of reasons to talk of causes (e.g. in terms of neurological processes). Dennett is however not as convinced of the success such an approach will yield, as he seems to believe that talk of reasons (folk-psychology in this case) has its own merits and is definitely more efficient: it is quite hard to understand another human being in terms of the complete neurological processes, whereas it is relatively easy to make sense of another human being in terms of beliefs, desires and other mental states.

One of the important points in this thesis has been that one should not think of the meaning of a concept in an essentialistic way. A good example of concepts that cannot be explained in an essentialistic manner, are metaphorical uses and secondary senses. The difference between sentences such as “ploughing through language” (metaphor) and “ploughing through earth” (literal) is an example of this. Another interesting subject to consider, is thus how the views on representation exposed in this thesis can help us to shed light on subjects such as metaphorical language use and secondary sense. This latter subject has been investigated by Ter Hark (e.g. Ter Hark, 1990) and is an interesting intermediate between literal and metaphorical language. When we for example call the pitch of a musical tone “high”, this is neither a literal nor a metaphorical use of the word.

CHAPTER 5 – CONCLUSIONS

Ter Hark suggests that the mastery of this secondary sense of the word is dependent on the mastery of the ‘normal’ use of the word in the first place (hence the term *secondary* sense, cf. PI, p. 216). That is, without the regular meaning of the word “high”, it would not be possible to call a pitch “high” (in the sense we are doing that now). Moreover, certain aspects of this phenomenon have been related to synaesthesia, or light versions of it, as some people have a consistent ‘experience’ of e.g. colour when certain words are being used. Wittgenstein himself gives the examples of “lean Tuesday” and “fat Wednesday” and also discusses an example that is nearer to synaesthesia: “If I say “for me the vowel *e* is yellow” I do not mean: ‘yellow’ in a metaphorical sense,-for I could not express what I want to say in any other way than by means of the idea ‘yellow’.” (PI, p. 216). These issues are discussed, amongst other places, in Part II of the *Philosophical Investigations*, which contains a treasure of valuable remarks still to be investigated more thoroughly by commentators.

This brings me to the next possible continuation of this research: in Part II a number of comments on aspect perception can be found. The phenomenon of aspect perception is about the ‘switch’ in perception you experience with figures such as the Necker-cube or the infamous duck-rabbit picture. The puzzling nature of this phenomenon resides in the fact that when looking at such a figure, one sees that something is different, while at the same time one also sees that the figure has not changed. After having investigated this to some extent, I have come to believe that again the solution to this problem can be traced back to an important type of conceptual confusion that Wittgenstein wishes to combat: the idea that we can explain all cases of a concept (“seeing”) by one particular occurrence or feature of it. Part of the explanation Wittgenstein offers for this phenomenon, is that being able to experience such a change of aspect is only possible after having mastered certain other skills. For example, being able to see the duck-rabbit picture now as a duck and later as a rabbit, is only possible if we can see a regular, unambiguous duck-figure as a duck and a regular rabbit picture as a rabbit. Again we see Wittgenstein’s emphasis on training, learning, activities and practices that we focused on in the final part of Chapter 4.

An important facet of section xi of Part II, where both the remarks on aspect perception and secondary sense are placed, seems to be that Wittgenstein shows a way in which we can establish certainty in our form of life: in training and learning to apply certain concepts, we also learn to take things in as a basic certainty, as an unquestionable basis. Learning to employ concepts such as ‘red’, ‘apple’ and ‘pain’ is related to basic activities we perform, ones that establish certainty and creating a basic and unquestionable ground: our form of life. All judgements that a community shares are in this unquestionable ground, without it, one cannot understand the form of life completely or become a full member of it. Learning this form of life means learning a way to judge,

CHAPTER 5 – CONCLUSIONS

learning a way to live: it is therefore mainly a psychological and mental process, and not ‘just’ a matter of behavioural training – which would invite a more causal explanation.

These considerations lead naturally into an investigation of Wittgenstein’s later work, of which we can find many remarks in *On Certainty*. I feel that this work has a lot of potential for further explorations, and it has many connections with Wittgenstein’s other considerations, such as *Cause and Effect*, *Remarks on Frazer’s Golden Bough* (to name only a few), but also, interestingly enough, Wittgenstein’s first work, the *Tractatus Logico-Philosophicus*. In it, I think, an early draft of Wittgenstein’s view on language can be found that has been attributed mostly to his later philosophy. The differences between the *Tractatus* and Wittgenstein’s later work might turn out to be not as substantial as is currently held by the majority of the commentators. But that is something to be considered at another occasion.

Bibliography

Abbreviations

- BB Wittgenstein, L. (1958), *The Blue and Brown Books*, Oxford: Blackwell.
- BT Wittgenstein, L., *The Big Typescript* [TS 213].
- CE Wittgenstein, L. (2003), 'Cause and Effect: Intuitive Awareness' in: *Human Knowledge, Classical and Contemporary Approaches*, Paul K Moser & Arnold vander Nat (eds.), Oxford: Oxford University Press. (Originally published as 'Cause and effect: intuitive awareness' in *Philosophia*, 1976/6, Edited by Rush Rhees, translated by Peter Winch, pp. 392-445.)
- LC Wittgenstein, L. (1978), *Lectures and Conversations on Aesthetics, Psychology and Religious Belief*, Oxford: Blackwell.
- PI Wittgenstein, L. (1953). *Philosophical Investigations*. G.E.M. Anscombe and R. Rhees (eds.), Oxford: Blackwell.
- RC Wittgenstein, L. (1980), *Remarks on Colour*, Oxford: Blackwell.
- RF Wittgenstein, L. (1979), "Remarks on Frazer's 'Golden Bough'", in C. Luckhardt (ed.), *Wittgenstein. Sources and Perspectives*, pp 61–81, Ithaca, N.Y.: Cornell University Press.
- RFM Wittgenstein, L. (1937-44), *Remarks on the Foundations of Mathematics*, revd edn, G.J. von Wright, R. Rhees, G.E.M. Anscombe (eds.), Oxford: Blackwell.
- TLP Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*, trans. D.F.Pears & B.F.McGuinness. London: Routledge.
- WLFM Wittgenstein, L. (1976). *Wittgenstein's lectures on the Foundations of Mathematics, Cambridge 1939*, notes taken by R.G.Bosanquet, N. Malcolm, R.Rhees and Y.Smithies, C. Diamond (ed.), Hassocks: Harvester Press.

Further References

- Ammereller, E. (2001), 'Wittgenstein on Intentionality', in *Wittgenstein: A Critical Reader*, Glock, H.J. (ed.), Oxford: Blackwell.
- Baker, G.P. & Hacker, P.M.S. (1980). *An Analytical Commentary on Wittgenstein's Philosophical Investigations*, volume 1. Oxford: Blackwell.

- Bennett, M.R. & Hacker, P.M.S. (2003) *Philosophical Foundations of Neuroscience*, Oxford: Blackwell.
- Bouveresse, J. (1995), *Wittgenstein reads Freud: The Myth of the Unconscious*, Princeton: University Press.
- Crane, T. (1995). *The Mechanical Mind: A Philosophical Introduction to Minds, Machines and Mental Representation*, Penguin USA.
- Crary, A. & Read, R. (2001), *The New Wittgenstein*, London: Routledge.
- Dennett, D.C. (1987), *The Intentional Stance*, Cambridge, Mass.: MIT Press.
- Dennett, D.C. (1991a), *Consciousness explained*, Boston: Little Brown.
- Dennett, D.C. (1991b), "Real Patterns" in *The Journal of Philosophy*, 89, pp. 27-51.
- Dennett, D.C. (1994), "Dennett, Daniel C." in Guttenplan, S. (ed.), *A Companion to the Philosophy of Mind*, Oxford: Blackwell.
- Dennett, D.C. (1996), *Kinds of Minds*, London: Weidenfeld & Nicolson, 1996.
- Dretske, F. (1980), "The intentionality of cognitive states," in Rosenthal, D. (ed.) *The Nature of Mind*, Oxford: Oxford University Press.
- Dretske, F. (1988), *Explaining Behaviour: Reasons in a World of Causes*, Cambridge, MA: MIT Press.
- Dretkse, F. (1994), "Dretkse, Fred" in Guttenplan, S. (ed.), *A Companion to the Philosophy of Mind*, Oxford: Blackwell.
- Fodor, J.A. (1975), *The Language of Thought*, Harvard: Harvard University Press.
- Glock, H.J. (1996), *A Wittgenstein Dictionary*, Oxford: Blackwell.
- Grice, H.P. (1957), 'Meaning' in *Philosophical Review*, 66, pp. 377-88.
- Hacker, P.M.S. (1990), *Wittgenstein: Meaning and Mind. An analytical commentary on the Philosophical Investigations*, volume 3. Oxford: Blackwell.
- Hacker, P.M.S. (2000), 'Was he trying to whistle it?' in Crary, A. & Read, R. (eds.), *The New Wittgenstein*, London: Routledge.
- Hacker, P.M.S. (2001), 'Philosophy', in: *Wittgenstein: A Critical Reader*, Glock, H.J. (ed.), Oxford: Blackwell.
- McGinn, M. (1997), *Wittgenstein and the Philosophical Investigations*, London: Routledge.
- Ogden, C.K. & Richards, I.A. (1923), *The Meaning of Meaning*, London: Harvest.
- Peijnenburg, J. (1998), "The Dual Character of Causality. How Davidson's Theory of Action Explanation Deals With Akrasia", *Dialektik* 35 (2), pp. 71–82.
- Putnam, H. (1974), "The meaning of 'meaning'," in Putnam, H. (1975) *Philosophical Papers*, vol. II, *Language, Mind and Reality*, Cambridge: Cambridge University Press.
- Putnam, H. (1992), *Renewing Philosophy*, London: Harvard University Press.
- Russell, B.A.W. (1921), *The Analysis of Mind*, London: Allen and Unwin.
- Savickey, B. (1999), *Wittgenstein's Art of Investigation*, London: Routledge.

- Slors, M.V.P. (1996), "Why Dennett cannot Explain what it is to Adopt the Intentional Stance", *The Philosophical Quarterly*, 46 (182), pp. 93-8
- Stroll, A. (2000), *Twentieth-Century Analytic Philosophy*, New York: Colombia University Press.
- Ter Hark, M.R.M. (1990), *Beyond the Inner and the Outer*, Dordrecht: Kluwer Academic Publishers.
- Thornton, T. (1998), *Wittgenstein on Language and Thought: The Philosophy of Content*, Edinburgh: Edinburgh University Press.
- Waismann, F. (1965), *The Principles of Linguistic Philosophy*, London: Macmillan.
- Waismann, F. (1979), *Ludwig Wittgenstein and the Vienna Circle: Conversations Recorded by Friedrich Waismann*, McGuinness, B.F. (ed.), Oxford: Blackwell.
- Waismann, F. (2003), *The Voices of Wittgenstein: The Vienna Circle: Ludwig Wittgenstein and Friedrich Waismann*, Baker, G. (ed.), London: Routledge.