Stockholm University

Measuring Health

On the Theoretical Foundations of Health Status Evaluations

Amanda Thorell

Doctoral Thesis in Philosophy at Stockholm University, Sweden 2021

Measuring Health

On the Theoretical Foundations of Health Status Evaluations

Amanda Thorell

Academic dissertation for the Degree of Doctor of Philosophy in Philosophy at Stockholm University to be publicly defended on Wednesday 8 September 2021 at 13.00 in hörsal 2, hus A, Universitetsvägen 10 A.

Abstract

This thesis is about the notions of health and pathology in medical theory. I develop a theory, which defines 'health' and 'pathology' in a way that solves several problems with earlier suggestions of how to define these terms. I call the theory 'the disposition profile efficiency theory', abbreviated 'the DPE-theory'. According to the DPE-theory, a trait token (e.g. an organ) is healthy, roughly, if and only if all of its dispositions for performing physiological functions are efficient enough. A trait token is pathological, roughly, if and only if at least one of its dispositions for performing a physiological function is not efficient enough. The notion of efficiency, I suggest, is reference class-relative: the efficiency of a trait token's disposition for performing a physiological function expresses a relation between the trait token and a health standard for the trait token's bearer's reference class.

The thesis also examines the most discussed theory of health and pathology, "the biostatistical theory" proposed by Boorse. Both the DPE-theory and the biostatistical theory are evaluated against a number of desiderata: the theory should (i) be theoretically sound; (ii) only use empirical, statistical, and logical terms; (iii) not involve values; (iv) be clear; (v) both account for health and pathology as reference class-relative properties, and account for the importance of integration of different physiological functions for health; and (vi) have reasonable implications. It is argued that the DPE-theory satisfies the desiderata, and that it does so better than the biostatistical theory.

The main contributions of the thesis are the DPE-theory's models of dispositions, its approach to defining reference classes, its efficiency measure, and its way of drawing the line between high enough efficiencies and too low efficiencies. Other important contributions concern the desideratum about reasonable implications. It is shown how the DPE-theory contributes to solving two much discussed objections directed towards the biostatistical theory, "the problem of common diseases" and "Kingma's dilemma". Further, the DPE-theory is used to illuminate discussions about certain types of conditions for which it is contested whether they should count as healthy or pathological.

Keywords: *biostatistical theory, disease, disposition, efficiency, function, health, pathology, philosophy of medicine, physiology, reference class.*

Stockholm 2021
http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-194042

ISBN 978-91-7911-524-1
ISBN 978-91-7911-525-8

Stockholm University

Department of Philosophy

Stockholm University, 106 91 Stockholm

MEASURING HEALTH

Amanda Thorell

Stockholm
University

Measuring Health

On the Theoretical Foundations of Health Status Evaluations

Amanda Thorell

©Amanda Thorell, Stockholm University 2021

ISBN print 978-91-7911-524-1
ISBN PDF 978-91-7911-525-8

Cover photo: Lars Thorell

Printed in Sweden by Universitetsservice US-AB, Stockholm 2021

Acknowledgements

In one sense, writing this thesis has been a one-person job. In another sense, it has been a job for many people: the result would not have been possible without the feedback and support I have got from various persons.

First of all, I want to thank my supervisors Sören Häggqvist and Richard Dawid for reading and commenting on all of my drafts. In particular, I want to thank Sören for his careful and often very detailed comments, and Richard for his thoughtful comments about the thesis as a whole. I also want to thank Miira Tuominen and Anandi Hattiangadi for reading and commenting on a late version of the thesis.

The thesis has also benefitted a lot from discussions with other people. I especially want to thank Karl Nygren for being a great discussion partner and observant reader. Also, both Karl Nygren and Anders Lundstedt deserve special acknowledgement for sharing their expertise in logic.

What have also been important are of course nice colleagues. I especially want to thank Lisa Hecht for support and friendship. I also want to mention Tara Nanavazadeh, who has been a kind officemate. Thanks also to Mikael Janvid, Åsa Burman, Daniel Ramöller, and Olle Torpman, who are always friendly and supportive.

Further, I want to thank everyone else at the Department of Philosophy at Stockholm University, including all PhD-students, administrative staff, and others. Thanks also to everyone at the Stockholm Center for Health Care Ethics.

Contents

1 Introduction **1**

1.1 Positioning My Theory in the Health and Disease Debate 3

1.1.1 Motivation 4

1.1.2 Concept Domain and Language Usage 6

1.1.3 Term and Extension 9

1.1.4 Scope 10

1.1.5 Organisms or Body Parts? 11

1.1.6 Individual- or Reference Class-Relative Properties? 12

1.1.7 The Relation between Health and Disease 16

1.1.8 Absolute or Comparative? 17

1.1.9 Is Disease a Natural Property? 17

1.1.10 Value-Involvement 18

1.2 Desiderata 19

1.2.1 Theoretical Soundness 19

1.2.2 Only Empirical, Statistical, and Logical Terms 20

1.2.3 No Value-Involvement 20

1.2.4 Clarity 20

1.2.5 Reference Class-Relativity and Integration 20

1.2.6 Reasonable Implications 21

1.3 Outline of the Thesis 23

2 The Biostatistical Theory **25**

2.1 Aim 25

2.2 Content 27

2.2.1 Reference Classes 28

2.2.2 Physiological Functions 30

2.2.3 Efficiency 37

2.2.4 Statistical Normality 39

2.3 Objections 42

2.3.1 Individuation of Reference Classes 43

2.3.2 The Choice of Goals 48

2.3.3 Kingma's Dilemma 51

2.3.4 The Problem of Common Diseases 57
2.3.5 Further Objections 59

3 Introducing the Disposition Profile Efficiency Theory 61
3.1 Basic Assumptions 62
3.1.1 Relevant Goals 62
3.1.2 Measuring the Goals 66
3.1.3 Treating the Goals Separately 69
3.1.4 Reference Classes, Trait Types, and Physiological Functions 70
3.2 Dispositions to Perform Physiological Functions 72
3.2.1 Feature Types and Feature Values 72
3.2.2 Feature Value Configurations 74
3.2.3 Specific Disposition Profiles 74
3.2.4 Complete Disposition Profiles 76

4 Reference Classes and Standards for Health 79
4.1 A Reference Class-Relative Standard for Health 79
4.1.1 Hausman's Standard 80
4.1.2 Exemplary Complete Disposition Profiles 80
4.2 Reference Classes 85

5 Efficiency, Health, and Pathology 95
5.1 Efficiency 95
5.1.1 Efficiencies of Dispositions 96
5.1.2 Hausman's Account of Efficiency 96
5.1.3 Comparing Dispositions 100
5.1.4 A Measure of Efficiency 104
5.2 Health and Pathology 107
5.2.1 Premature Definitions 108
5.2.2 Final Definitions 110

6 The Problem of Common Diseases and Kingma's Dilemma 115
6.1 Solving the Problem of Common Diseases 115
6.1.1 The Statistical Fitness Theory 116
6.1.2 The Frequency and Negative Consequences Approach 119
6.1.3 The Functional Efficiency Theory 121
6.1.4 A Biostatistical Account of Functions 123
6.1.5 The Disposition Profile Efficiency Theory 125
6.2 Solving Kingma's Dilemma 127

6.2.1 Appeal to Biological Normality 128
6.2.2 Allow for Differently Fine-Grained Descriptions of Situations 129
6.2.3 Consider the Trait Token's Disposition to Perform the Function in Different Situations 130
6.2.4 The Disposition Profile Efficiency Theory 132

7 Contested Cases 137
7.1 Trivial Conditions 137
7.2 Normal Ageing 140
7.3 Defense Mechanisms 143
7.4 Risk Conditions 146
7.5 Summary 150

8 Conclusion 151

Appendix: Definitions 159

Svensk sammanfattning 163

References 181

1 Introduction

This thesis is about health and pathology, or disease. A very rough analysis, which in principle everyone would agree on, is that health is about functioning well enough and that pathology is about not functioning well enough. Disagreements in the debate regard how to understand "functioning" and how to understand "well enough".

The notions of function, health and disease have been discussed since Aristotle. The modern debate about these notions was intensified after 1973, when Wright broke new ground for the function debate with the following analysis (1973,161):

> The function of *X* is *Z means*
> (*a*) *X* is there because it does *Z*,
> (*b*) *Z* is a consequence (or result) of *X*'s being there.

This idea of how to understand functions started off an intense debate about functions in general, but also a debate about how to understand biological, or physiological, functions in particular.[1] This, in turn, fueled the debate about health and disease from a biological perspective: New ideas about physiological functions both provided immediate ideas about how to understand the function part in "not functioning well enough", and, indirectly, gave new input to the issue of how to understand "well enough". For example, on Boorse's (1976b) goal-analysis of physiological function, physiological functions are contributions to survival or reproduction. With this analysis, a natural way to cash out "well enough" is in terms of sufficient contribution to survival and reproduction.

The debate about health and disease that started off after Wright (1973) tended first to circle around the opposition between a biological aspect (concerning dysfunction) and a normative aspect (concerning a negative valua-

[1] See e.g. Bigelow and Pargetter (1987), Boorse (1976b), Cummins (1975), Godfrey-Smith (1993), Kitcher (1993), Millikan (1984), Neander (1991).

tion). According to Boorse (1977), one of the pioneers in the health and disease debate, 'disease' can be defined as a theoretical biological concept without the involvement of any normative component. According to many others (e.g. Agich 1983; DeVito 2000; Engelhardt 1975; Hare 1986; Reznek 1987; Wakefield 1992), the disease concept inevitably involves some value component. Although the debate today is more multifaceted, the opposition between a purely biological, value-free view on the one hand, and a value-based or value-involving view on the other, is still an ongoing theme (Kingma 2017).

The thesis investigates the view of disease as a theoretical entity in biology, or, more narrowly, in physiology and pathology. A basic assumption made in the thesis is that natural and normative considerations can be separated from each other. A special focus is given the notion of efficiency. This is an important component in Boorse's theory which, in contrast to other parts of his theory, has not been much discussed.[2] According to Boorse, functions may be performed with different efficiencies, and these differences are crucial for health status evaluations. For example, a heart pumping blood with typical efficiency is healthy whereas a heart pumping blood with untypically low efficiency is diseased. In Boorse's theory, however, the notion of efficiency is left more or less undefined. This is problematic, since in order to evaluate the view of disease as a theoretical entity in physiology and pathology, and to specifically address some of the worries directed towards Boorse's theory, the notion of efficiency must be fully accounted for. Otherwise, it is neither clear whether the theory actually works out in detail, nor what the theory implies about particular cases.

In this thesis, I develop an alternative theory of health and disease. When developing this theory, I put a sharp focus on the notion of efficiency. But I also try to be more precise than Boorse in all parts of the theory. The theory that I suggest solves some serious problems posed against Boorse's theory, and explains better why some objections posed against it are mistaken. It also enlightens discussions about controversial cases, i.e. where intuitions about health and pathology tend to go in opposite directions. Examples of this concern dysfunctional states that are trivial from the perspective of the whole organism (e.g. a single malfunctioning cell), normal ageing (e.g. presbyopia, diminished muscular strength, reduced memory capacity), defense mechanisms (e.g. fever, vomiting, diarrhea), and risk conditions (e.g. hypertension, high cholesterol, osteoporosis, obesity).

[2] An exception here is Hausman (2012).

The philosophical debate about how to define ‘health’ and ‘disease’ and similar notions has become rather extensive since Boorse first presented his theory in 1977. The debate is also sprawling; theorists within the debate discuss the notions from different angles, and for different reasons. In order for this thesis to make a meaningful contribution to the debate, its aims and basic assumptions should be clarified in relation to the already existing landscape of theories. I will therefore start by describing the philosophical health and disease debate, and, adjacently, position the research of the thesis in relation to it (Section 1.1). I will also provide a list of desiderata which will serve as a basis for evaluation throughout the thesis (Section 1.2). Last in this chapter, I will give an overview of the seven remaining chapters (Section 1.3).

1.1 Positioning My Theory in the Health and Disease Debate

Theorists within the philosophical debate about health and disease differ on a range of parameters, with regard both to their aim, their methodology, and their theories’ content. Unfortunately, it is generally the case that some key terms for describing and classifying theories have been used routinely without giving them clear definitions. Here, I am especially thinking about the terms ‘naturalism’ and ‘normativism’. These terms are often implicitly taken to include several, logically independent, views. This has resulted in a somewhat muddy debate. In recent years, though, contributions have been made in order to elucidate the discussion (see e.g. Broadbent 2017; Lemoine and Giroux 2016; Simon 2007). I will continue in the same spirit, spelling out clearly the meanings of the terms I use for different positions. To further evade confusion, I will as far as possible avoid using the original muddled terms.

Beneath, I will map out the philosophical debate about health and disease and position the theory that will be developed in this thesis within this map. When doing this, I am more interested in the logical landscape of viewpoints about health and disease and less interested in organizing and categorizing all various existing theories of health and disease. The map takes the form of ten questions. I will discuss how one may answer these and note how possible answers to certain questions may affect how one reasonably answers, or may tend to answer, some of the other questions.

1.1.1 Motivation

The most central question is what motivates defining 'health' and 'disease'. This does, as we will see, affect what are reasonable answers to more or less all of the other nine questions.

The majority of theorists are motivated by ethical concerns of some sort, for instance political or legal issues, issues of health care allocation, or issues about everyday behavior. Let us call their motivation 'practical'. The common thought among these theorists is that definitions of 'health' and 'disease' will help us answer difficult questions, for example:

- Who should be entitled to public health care?
- Who should be entitled to sickness benefits?
- When is it legitimate to medically treat a person without her consent?
- When should a person be held responsible for her actions?
- When should a person be given medical care rather than a prison penalty?
- Which undesirable behaviors are excusable?
- Which conditions are such that society should take extra responsibility for those having them?
- Which activities should be sanctioned as health promoting?

Another sort of motivation for theorizing about health and disease is theoretical. Some theorists are interested in the concept of disease in medical theory. Here, the idea is that definitions of 'health' and 'disease' will help us conceptually strengthen, or better understand, medical theory, which is typically taken to consist of physiology and pathology. The theoretical motivation and the practical motivation are not mutually exclusive. They may be adopted simultaneously, although this might make it trickier to come up with satisfying definitions.

It has been argued that a purely theoretically oriented project is not interesting or adequate. For example, Nordenfelt writes that "[a] theory of health should be able to perform many tasks, especially tasks that have to do with the care and treatment of human illness" (2018, 12). Similarly, Worrall and Worrall complain about Boorse's purely theoretically motivated theory that

> […] it is difficult to see it of major interest: practicing clinicians wanted a clear-cut notion of disease to enable them to give clear-cut answers to "hard-headed

> lawyers" and others, but Boorse admits that his own analysis, even if successful, would not meet this need. (2001, 45)

From a practically oriented perspective, it is not strange to consider theoretically motivated theories uninteresting. However, purely theoretically motivated theories should not be ruled out as plainly uninteresting on the grounds that they to not answer the interests given by a practical motivation. Whether purely theoretically motivated theories are interesting or not must depend on the role the concepts of health and pathology play in medical theory. I will discuss this role beneath.

Some philosophers reject both practical and theoretical motivations for defining 'health' and 'disease', arguing that there is no benefit in defining 'health' and 'disease' whatsoever. Hesslow (1993) calls the disease concept 'a straitjacket'. A definition of 'disease', he argues, is not the best help for answering practical questions. For example, medical treatment should not be justified by the presence of disease *per se*, but by potential benefits of treatment (Hesslow 1993, 7). And special rights granted to people with a disease (e.g. the right not to work) should not be justified by the presence of a disease, but rather by considerations about discomfort and risk (Hesslow 1993, 9). A definition of 'disease' will not, Hesslow argues, contribute to medical theory either. This, he thinks, is because 'disease' is neither a theoretical term in the same strong sense as 'electron' or 'gene', nor a theoretical term in the weaker sense of merely playing a role in medical and biological thinking.

A potential role of the disease concept in medical and biological thinking, which Hesslow discusses, is to help define the area of interest for physiology. But the disease concept does not play this role, Hesslow argues, because physiologists' interest is not limited to describing healthy organisms. Rather, Hesslow writes,

> Physiologists take it as their main business to study anything that is physiologically important. The goal of this research is to produce a body of knowledge that can explain, predict, and control important physiological phenomena, regardless of the reasons for this importance. Usually, this means normal or at least common physiological phenomena, but that is only because scientific importance is partly of [a?] function of the range of applicability of a certain finding. (1993, 11)

Similarly to Hesslow, Ereshefsky argues that we should give up "trying to find the correct definitions of 'health' and 'disease'" (2009, 221). Instead, when "discussing controversial medical issues", Ereshefsky suggests, we should

"explicitly talk about the considerations that are central in medical discussions, namely, state descriptions and normative claims" (2009, 225).

Hesslow and Ereshefsky make valuable points. Concerning practical questions, I think they are right that it is more reasonable to try to answer these with reference to value claims about physiological states (or these states' consequences), rather than with reference to definitions of 'health' and 'disease'. Concerning their arguments that there is no point in defining 'health' and 'disease' for theoretical purposes, though, I do not fully agree. Hesslow may well be right that 'disease' is not a theoretical term in the same strong sense as 'electron' or 'gene'. But is there no other role that the disease concept may play in biological and medical thinking other than defining the range of interest for physiology?

The theory of health and disease that I will develop in this thesis is purely theoretically motivated. And given the definitions I will suggest, there is a potential role for these concepts to play in medical theory. According to the theory that will be developed, the concepts of health and disease, and their constituent concepts, may play the role of communicating knowledge about the connection between physiological states and different goals of interest (e.g. survival or reproduction), in relation to types of organisms. This idea of health and disease is more complex than only involving token state descriptions (or normative claims). It involves reference to specific goals, to types or reference classes of organisms, and to an account of what "is normal for" or "could be expected of" tokens of different body part types in a reference class. These relations may be considered physiologically interesting by themselves. They offer help in understanding the biological world by means of categorization.

This said, Hesslow and Ereshefsky may still be right that the notions of health and disease are not as central in medical theory as they may have appeared, at least in the philosophical debate. Yet, investigating how 'health' and 'disease' may be defined as theoretical terms is a way to make the discussion of the potential benefit of the terms in science better informed.

1.1.2 Concept Domain and Language Usage

Most theorists assume that there is only one concept of disease to analyze, or at least only one that makes sense to analyze (e.g. Brülde and Tengland 2003; Nordenfelt 2018). That concept, they think, is a common, or shared, concept of disease. Other theorists distinguish between different disease concepts.

Boorse (1977) defines a theoretical concept of disease, which he distinguishes, for example, from social and clinical disease concepts.

The domain of interest is naturally connected to one's motivation for defining 'disease'. If one has a broad practical motivation, it may be more convenient to think of the disease concept as a common concept. For example, Nordenfelt defends the choice to analyze the concept of disease as a broadly shared concept:

> […] it is this *ordinary* notion of health that defines the *aim* of many human activities, and in particular the aim of health enhancement and health care. What is crucial to us human beings is our well-being and our ability to do important things in life. […] We are normally not primarily concerned about the fact that our organs give a statistically normal contribution to survival. The whole rationale of the clinic lies instead, I would say, in the ordinary lay notion of health. (2018, 12, italics are original)

If one instead, like Boorse, is motivated by theoretical concerns in physiology and pathology, then it is convenient to distinguish the concept of disease in physiology and pathology from other possible disease concepts. Since my interest is connected solely to the theoretical motivation, this thesis will be concerned with a purely theoretical notion of disease.

Whether physiology and pathology are the only biological or medical disciplines employing the notions of health and disease has been questioned. Lemoine and Giroux (2016) suggest that the notions of health and disease may also, for example, be part of epidemiology, bacteriology, and genetics. To limit the aim of this thesis, the definitions of 'health' and 'disease' that I will suggest are mainly focused on physiology and pathology.

Regardless of what domain one takes the concept of disease to belong to, there is a question about how much one's definition must conform to the usage of 'disease' in that domain. For example, if one defines 'disease' as a theoretical term in physiology and pathology, this definition should reasonably conform at least to some extent to physiologists and pathologists' usage of 'disease'. And if one defines 'disease' as a legal term, this definition should reasonably conform to some extent to lawyers' usage of the term. And if one takes there to be only one concept of disease, common for pathologists, clinicians, lawyers and lay people, the definition of 'disease' should reasonably conform to some extent to all these people's usage of the term.

We may distinguish between two rough positions with regard to the question of degree of conformity to language usage. First, we have descriptive ac-

counts, which aim to describe current usage of the term. For a descriptive account to be successful, it must have a very high degree of conformity to language usage in the relevant domain. This sort of account is often accompanied by the method of conceptual analysis, where the aim is to discover, rather than to invent, the meaning of the term in question.

In contrast to descriptive accounts, there are revisionary or explicative accounts where the importance of conformity to language usage is curtailed in favor of other assumed virtues. Revisionary and explicative accounts may also be described as accounts resulting from conceptual engineering.[3]

Considering theories about non-scientific notions of health and disease, revisionist accounts suggest revisions of the usage of 'disease', for example for political purposes. The investigation may for instance be led by the question what definitions of 'health' and 'disease' will be the most helpful in organizing a fair society. In contrast to a descriptive definition, a revisionist definition may be seen more as an invention than a discovery. And a successful account need not conform much to the relevant language usage. Rather, what matters is whether the definition fulfills one's purposes for defining 'disease'.

Considering theories about scientific notions of health and disease, one may also talk about revisionary accounts. However, here it is more common to contrast descriptive accounts with explicative accounts. Explicatory accounts aim at improving, rather than describing, medical theory. For an explicative account of disease to be successful, it should stay as close as possible to language usage in medical theory, but only insofar as 'disease' is given a clear and consistent definition, and insofar as the role of the disease concept comes out as theoretically fruitful. These requirements may in principle result in explicative definitions that depart from language usage in medical theory. Also, an explicative aim does not preclude several different definitions of 'health' and 'disease', given that this is required in order for medical theory to be consistent or in order to maximize theoretical fruitfulness.[4]

What sort of definition one aims at is related to one's motivation for defining 'disease', in combination with one's thoughts about the status of the usage of the term in the relevant domain. For example, if one is practically motivated by the question of which states should be treated, and one also believes that the right answer to this question is already embedded in our usage of 'disease',

[3] The debate about conceptual engineering comprises different ideas about what desiderata there are for this method. See e.g. Plunkett and Cappelen (2020).

[4] For more discussion about explicative accounts of disease (especially in comparison to conceptual analysis), see Schwartz (2014) and Varga (2020).

then one reasonably aims at a descriptive definition of 'disease'. If one is practically motivated by the question of which states should be treated, but instead believes that our common usage of 'disease' is not good guidance for answering this question, then one reasonably aims at a revisionist definition of 'disease'.

If one is theoretically motivated, one may either be interested in merely describing current usage of 'disease' in medical theory, and thereby aim at a descriptive account. Or, if one considers current usage of 'disease' in this domain to be contradictory or theoretically fruitless, one may instead aim at an explicative account.

The account that I will develop is explicative. This choice is related to my motivation. As already indicated, I want to define 'health' and 'disease' as terms in medical theory. Looking at the vast philosophical literature on the topic, where the issue of defining 'health' and 'disease' is still unresolved, it seems that formulating a successful descriptive account is not possible. Also, if we consider Boorse's theory, several objections stem from a view of this theory as descriptive rather than explicative.[5] The most promising option is, hence, to go for an explicative account. This choice should also be considered more interesting than a descriptive account. An explicative project will not just examine how 'health' and 'pathology' are used, but whether 'health' and 'pathology' may be defined in a way that makes them clear, coherent and fruitful concepts in medical theory. When developing my theory, I will use very basic ideas about health and disease in medical theory as guidelines for the theory[6], but I will also aim for a theory answering to the theoretical virtues of coherence, clarity, and theoretical fruitfulness.

1.1.3 Term and Extension

So far, I have talked about theories of disease. 'Disease' is, however, not the term that everyone within the debate chooses to define. Whereas Boorse

[5] For example, Lemoine and Giroux (2016) argue that Boorse's theory is not a faithful description of physiology and therefore fails as a descriptive attempt to define 'health' and 'disease'. Doust, Walker and Rogers (2017) question whether Boorse's usage of reference classes always fits the reference classes employed in medical theory. Scadding (1988) thinks that Boorse's theory conflicts with current nosology, and Nordenfelt (2001) points out that some entries in ICD-10 involve no physical dysfunction, which for Boorse is a necessary requirement for disease. Both Nordenfelt (2001) and Stempsey (2000) point to pathology textbook definitions of 'pathology' that differ from the Boorse's definition. Varga (2020) argues that a descriptive definition of 'disease' cannot succeed because pathologists are not authorative about the concept of disease as Boorse supposes.

[6] See Sections 1.2.5 and 1.2.6 below.

(1977) and Cooper (2005) use ‘disease’, Fulford (1989) use ‘illness’, Wakefield (1992) use ‘disorder’, Culver and Gert (1982) and Nordenfelt (1987) use ‘malady’, and Boorse (2014) use ‘pathology’.

What one takes to be the relevant term may be influenced by the domain one takes the relevant concept to belong to. For example, if one exclusively aims at the concept in physiology and pathology, one should reasonably (at least if one also aims at a descriptive definition) choose a term that is commonly used for this concept among physiologists and pathologists. One may also, if possible, choose a term that is not as commonly used in other domains. In the following discussion, I will stick to the term ‘disease’, since this term is most commonly used within the debate. However, when later formulating my own theory, I will use the term ‘pathology’, since this term is used in physiology and pathology, but typically not among everyday people.

Although theorists disagree about which term to define, the very rough idea of the extension that is to be captured by the term is more or less common ground. The great majority aims at defining a term that refers to all conditions that are not “healthy” or “normal”. This means that they, for example, aim to cover infections as well as broken bones and congenital defects.

However, even if there is agreement at a very rough level about the extension of ‘disease’, there are controversies about the more precise extension. First, theorists differ in what they think is the scope of the disease concept – that is, to what areas it is applicable. I will discuss this issue next. Second, there are different views about exactly which conditions, given a certain area of application, should fall within the disease category. Some types of conditions whose health status theorists tend to disagree about concern (i) conditions that are trivial from the perspective of the whole organism (e.g. a single malfunctioning cell), (ii) normal ageing (e.g. presbyopia, diminished muscular strength, reduced memory capacity), (iii) defense mechanisms (e.g. fever, vomiting, diarrhea), and (v) risk conditions (e.g. hypertension, high cholesterol, osteoporosis, obesity). In the last chapter of the thesis I will discuss what the theory that I will develop implies about these different types of conditions.

1.1.4 Scope

There are a couple of questions about the scope of the disease concept, on which theorists take different standpoints. One question concerns what sorts of entities that can be healthy or have a disease. I do not know of any theorist who takes other entities than organisms (or parts of organisms) to be healthy

or diseased. However, theorists differ in which organisms they take the concepts of health and disease to apply to. Some theorists explicitly aim at covering all organisms (e.g. Boorse 1977; Hausman 2012). This is natural if one aims at defining 'disease' for purely theoretical reasons, since biology does not treat humans as ontologically different from other organisms. In line with most theoretically motivated theories, the theory that I will formulate is an account of disease common for all organisms.

Others provide theories that are explicitly supposed to not apply to, or do not seem possible to apply to, all organisms. For instance, some accounts of disease include a subjective harm requirement (e.g. Reznek 1987), which seems problematic to apply to lower organisms and plants. This scope may be reasonable if one is practically motivated by the question who should be entitled to public health care, since one then may want to limit the area of application to humans.

Another question of scope concerns the somatic-psychiatric distinction. Are somatic disease and psychiatric disease basically the same thing? Some think so, and suggest definitions that are meant to encompass both somatic and psychiatric disease (e.g. Cooper 2005; Fulford 1989). If one is practically motivated by the question who should be entitled to public health care, this may be a reasonable standpoint. Others, in contrast, think that there is an important difference between somatic and psychiatric disease, and that one therefore cannot encompass both by the same definition (e.g. Szasz 1974).

Some theorists account for somatic disease, and then leave it an open question whether the theory also applies to psychiatric disease (e.g. Boorse, 2014). This may be a sensible choice if one wants to avoid having to defend the physical or ontological status of psychiatric disease. This is also the strategy I will take when developing my theory: I will define somatic disease, but I will not deny that the definition I propose is common for somatic and psychiatric disease.

1.1.5 Organisms or Body Parts?

Theorists disagree about whether it is primarily whole organisms, or parts of organisms (e.g. cells, organs, systems), that are diseased or healthy.

If one takes the concepts of health and disease to primarily apply to whole organisms, this means that when considering an organism's health status, it does not directly matter what goes on, for example, at the level of the cell or organ. Rather, what matters are the gross effects at the level of the whole organism.

The alternative is to hold that the concepts of health and disease primarily apply to parts of organisms. If one is theoretically motivated and takes the disease concept one analyzes to belong to physiology and pathology, this, I think, is a reasonable view to take. Physiologists and pathologists are interested in understanding how organisms function. This means that they consider, not only whole organisms, but also the workings and interrelations between different parts of an organism. In order to be able to explain in detail why an organism is not functioning well enough, one needs to disassemble the organism into parts. Even if one wants to say of a whole organism that it is diseased, the organism is so in virtue of a specific part being diseased. Therefore, the account of health and disease that I will develop primarily ascribes health and disease to organism parts. We may note, though, that this approach does not preclude that an organism as a whole may be healthy or diseased in some derivative sense – for instance by having, or not having, only healthy parts.

1.1.6 Individual- or Reference Class-Relative Properties?

A basic thought among theories of disease is that when something is diseased, it does not work as well, or is worse off in some respect, than some standard of comparison. What does this standard consist in? Some explicitly take the standard to be relative to the concerned individual. For example, Nordenfelt takes the standard for health to be determined by the particular individual's "vital goals", which are defined as "the state of affairs that are necessary and jointly sufficient for his or her [the concerned individual's] minimal long-term happiness" (2018, 14). Alternatively, one may take the standard to be relative to the concerned individual's reference class.[7] On the first view, two exactly similar human heart tokens may have different health statuses because they are part of two different individuals of the same reference class. On the second view, if the two heart tokens are exactly similar, they must have the same health status (given that the two individuals belong to the same reference class).

The individual-relative view may be tempting to adopt. First, if one is practically motivated, this view may be motivated by the fact that when reasoning about interventions in medical practice, it is important to take into consideration the individual's preferences and needs. For example, Amundson, embracing an individual-relative concept of health, writes:

[7] The idea here is that reference classes comprise more than just one individual; i.e., one individual cannot make up its own reference class.

> Rather than testing a patient's statistical conformity, medical judgments should assess 'that congruence between physiological performance of the individual and the performance necessary for him. . . . An individual may be healthy–responsive–without regard to the quality or quantity of the morphology and function which the statistical norm would wish to prescribe for him' (Vácha, 1978 p. 826). Notice that the concept of responsiveness (individual normality) abandons the statistical and comparative basis of normality, replacing it with an assessment of the relation between individual performance and needs. There is no need for a species design. (2000, 44)

There also seems to be a more theoretically oriented pressing problem about the idea of health and disease being reference class-relative. That is the fact that, at a detailed level, in principle all organisms are unique. This fact means that when determining the health status of a token organ we cannot just compare its performance against some common standard for its organ type. Rather, we must see to what goes on in the rest of the organism.

Let us look at an illuminating example, somewhat simplified: The system regulating the human body's overall use of energy includes the thyroid, the pituitary and the hypothalamus. The regulation is done with two hormones – T3 (triiodothyronine) and T4 (thyroxine). These are secreted into the blood by the thyroid. Through the blood stream these hormones reach and enter cells of the body, in which they activate or repress the transcription of certain genes, and thereby regulate energy consumption. Whereas the thyroid produces T3 and T4, the pituitary and the hypothalamus regulate the amount of T3 and T4 which the thyroid produces and secretes. The pituitary secretes the hormone TSH (thyroid stimulating hormone, or thyrotropin). It is in response to this hormone that the thyroid produces and secretes T3 and T4. The pituitary, in turn, produces and secretes TSH as a response to the hormone TRH (thyrotropin releasing hormone), which is secreted by the hypothalamus. There is also a negative feedback loop. Both the hypothalamus and the pituitary are sensitive to circulating T3 and T4. The higher the levels of these hormones in the blood, the less TRH and TSH are secreted. The lower the levels of T3 and T4 in the blood, the more TRH and TSH are secreted.

The whole system described is rather complex. If we are to evaluate the health status of one part in the system regulating the human body's overall use of energy we cannot simply look at that part's performance in isolation and determine how well it works. Rather, we need to look at the whole system. For example, in order to evaluate the pituitary, we cannot merely look at what amount of TSH it secretes. We need to look at the amount of TSH it secretes

in relation to the amount of THR secreted by the hypothalamus and in relation to the amount of T3 and T4 in the blood stream.

This type of dependence between the performance of different physiological functions may be labelled 'integration'. I take integration to involve both direct dependencies at the genetic level (i.e. where certain dispositions for different physiological functions are inherited together) and dependencies resulting from inherited tendencies of different parts of an organism to adapt to each other (and the environment) after conception or birth.

Amundson (2000) and Vácha (1985) have invoked the phenomenon of integration to argue that we cannot use the same health standard for all individuals of a reference class in health status evaluations. Since the amount of variation between (healthy) individuals is so great, the argument goes, health statuses must be relative to individuals. Amundson and Vácha conclude that health and disease is a matter of the relation between the individual's needs and her capabilities.

In contrast to Amundson and Vácha, I will argue that health and disease should be accounted for as reference class-relative properties, at least if one is theoretically motivated to define 'health' and 'disease' as terms in medical theory. When developing my theory of health and disease in later chapters I will show that it is, in fact, possible to account for the importance of integration for health within a reference class-relative view.

A first reason to think that health and disease should be accounted for as reference class-relative properties is because of usage of these concepts in medical theory. Often in medical theory, when it is claimed that a token organ is diseased, what is claimed is that the token organ does not meet the standard for its organ type, rather than that the token organ does not meet the individual organism's standard for that token organ. If we look at physiological theory, we find descriptions of the workings of healthy organs and systems for groups of similar individuals, rather than for individuals. And looking at pathological theory, more or less typical impairments of these *types* of organs and systems, resulting in disease, are described. If one held that health and disease are individual-relative properties, one would have to admit that what is presented as a healthy heart for a specific reference class in physiology may not always be a healthy heart – whether it is is relative to the individual organism.

Although there may be disagreement about whether the health status of a token organ should be relative to the specific token organism or not, I think it is at least clear that there are some intuitions pulling in the direction that the health status of a token organ in a certain reference class should be the same irrespective of which token organism it is part of. And although the terms

‘health’ and ‘disease’ may sometimes be used in an individual-relative sense, I think it may be useful to make distinctions between health and disease in a reference class-relative sense. For example, if an organism is able to fully compensate for a reduction of a certain token organ $o1$ by an atypically high capacity of another token organ $o2$, then $o1$ would not count as pathological on an individual-relative view of health; it would, however, on a reference class-relative view on health. Even if we would assume an individual-relative view on health and judge $o1$ healthy, it would still be interesting to know that a reason why $o1$ is not pathological is that $o2$ compensates (beyond what could “be expected”) for the reduced capacity of $o1$.

Second, the indication that health and disease should be accounted for as reference class-relative properties becomes stronger if we combine the above conclusion with a requirement of theoretical fruitfulness. If health and disease were about the relation between a goal, an individual and a part (e.g. an organ) of this individual (a relation we may refer to rather easily without the terms ‘health’ and ‘disease’), there would be no (existing) terms left to take care of the more complex relation between a goal, an individual, a part of this individual, and a reference class of this individual, which according to the above indications seem to occupy at least some interest in medical theory.

Third, it is hard to see how to formulate a reasonable health standard in the absence of a reference class. To see this, let us conceive of a token creature that is of its single kind, and is not like any other organism (i.e. it does not belong to any reference class). Although we may measure how well parts of this creature serve certain goals, there is no given standard to compare this with. Without a reference class there is no answer to the question what could be expected of a certain part of this creature. Although physiologists may imagine ways in which a certain part of this creature would function better for a certain goal, what they imagine (or could imagine) cannot reasonably set the bar for health. If that would be the case, we would get “too much” disease. It would imply that in principle all human organs are pathological. For example, the human eye could be better constructed, for instance more similar to the eye of an eagle, and the human digestive system could be better constructed, for instance by being able to digest cellulose.

Having settled on the view that health and disease are reference class-relative properties, the question is how to account for the reference class-relative standard in a way that takes into consideration the phenomena of integration.

When developing my theory in later chapters, I will explain how this can be achieved.[8]

1.1.7 The Relation between Health and Disease

Above, I have talked mostly about disease. What is health in relation to disease?

Many theorists take health and disease to be mutually exclusive. According to Boorse (1977), for example, an organism is healthy if it does not have any disease. Others, who also take health and disease to be mutually exclusive, take health to include more than the absence of disease. An oft-cited example of this view is the World Health Organization's definition of health: "Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (World Health Organization 1948). One reason to define 'health' as more than the absence of disease may be a practical motivation. If one thinks that a definition of 'health' will help us answer questions such as which activities should be sanctioned as health promoting, then it seems reasonable to include more activities than merely those that are associated with lowered risk of disease. One may want to include activities that are beneficial for high life quality, or "extra good health".

There are also theorists who do not take health and disease to be mutually exclusive, but rather allow an organism to be healthy and have a disease at the same time. An example of this is Nordenfelt (2018), whose definitions of 'health' and 'disease' I mentioned in relation to the previous question. Nordenfelt writes about his definition of 'disease' that:

> […] not all diseases actually compromise health in the holistic sense of being able to realize vital goals. Some maladies are aborted, i.e. disappear before they have influenced the person as a whole; others are latent; yet others are so trivial that they hardly affect the person's abilities and may therefore never be recognized by their bearer. […] [A person] may have a disease but still remain completely healthy […] (2018, 15)

We may note that, on Nordenfelt's account, the relation between health and disease is connected to the issue of whether health and disease apply to whole organisms or to their parts. According to Nordenfelt, an organism as a whole may be healthy even though a part of it is diseased.

[8] See especially Chapter 3, Section 3.2.

On the account that I will develop, an organ or other body part is healthy if it is not diseased.[9] This is reasonable given that my motivation is only theoretical.

1.1.8 Absolute or Comparative?

Can something be more or less healthy, or more or less a disease? Most theories draw an absolute line, where everything on the one side is plainly healthy and everything on the other side is plainly diseased. In contrast, one may think that there is no such absolute line, but that there are only degrees of health. For example, Schroeder (2013) drops the disease concept and argues that we should only talk about gradations of health. Similarly, Brülde and Tengland (2003, 36-37) argue that the concept of health is comparative, since health comes in degrees – from very poor health to very good health. In contrast to Schroeder, though, Brülde and Tengland (2003, 36-37) do not drop the disease concept, but take this concept to be absolute.

The issue whether the concepts of health and disease are absolute or comparative is interesting, but not one that I will focus on in this thesis. The account that I will present will be formulated as an absolute account, both of health and of disease. Nevertheless, it should be possible to develop a comparativist account based on the ideas I put forward in the thesis.

1.1.9 Is Disease a Natural Property?

An ontological question, which theorists tend to disagree about, is whether disease is a natural or a non-natural property. I take natural properties to be properties in nature that can be defined in empirical, statistical, and logical terms.

If one has a purely practical motivation for defining 'disease', it may not matter so much whether disease is a natural property or not. But if one is purely theoretically motivated, and tries to define 'disease' as a biological term, then one should presumably view disease as a natural property. Since the theory that I will present aims at a theoretical concept of disease, I will adopt the hypothesis that disease is a natural property.

Let us also note that there is the narrower question whether disease is a natural kind or not. Some take the project of "naturalistic accounts" of disease to be to carve nature at its joints (see e.g. Varga 2020). However, that disease

[9] However, it may in some cases be indeterminate whether an organ or other body part is healthy or diseased. See Chapter 5, Section 5.2.2.

is a natural property does not imply that it is a natural kind. It may, for example, merely be a relation between some other natural properties that we find interesting. I will not assume anything about whether disease is a natural kind or not.

1.1.10 Value-Involvement

One issue, which has been part of the center of attention in the disease debate, is whether the concept of disease is value-involving. Whereas a few writers (e.g. Boorse 1977) claim that it is not, the majority of writers claim that it is. However, the ways in which they take the concept to involve some value component differ.

Some think that the disease concept is value-involving in the sense that its meaning is partially evaluative, i.e. so that deeming a condition a disease implies a negative evaluation of it. For example, Clouser, Culver and Gert (1997) think that calling a condition a disease implies condemning it as a condition that should be avoided or ended. And according to Engelhardt, "the concept of disease acts not just to describe and explain, but also to enjoin to action. It indicates a state of affairs as undesirable and to be overcome" (1975, 127). To hold that the disease concept involves some value component in this way is a natural standpoint to take if one is practically motivated, for example by the question of who should be entitled to public health care. However, if one is theoretically motivated, and defines 'disease' as a purely theoretical term in medical theory, one should reasonably hold that the disease concept is not value-involving in the above sense. A theoretical concept in the natural sciences should not involve any values.

Others take the disease concept to involve some descriptive value judgement. Here, to deem a condition a disease does not impose a general negative evaluation of that condition. Rather, that the disease concept involves some descriptive value judgement means that the extension of 'disease' is wholly or partially determined by some people's value judgements. For example, Wakefield holds that a requirement for a condition to count as a disease is that it "causes some harm or deprivation of benefit to the person as judged by the standards of the person's culture" (1992, 385). To hold that the disease concept is value-involving in this way may for example be natural if one is practically motivated and believes that the subjects, or the culture, whose judgement is referred to will provide the right answers for those practical questions. However, if one is theoretically motivated, and aims at defining 'disease' as a purely theoretical term in physiology and pathology, it is reasonable to dismiss

the claim that the disease concept is value-involving in this sense. A theoretical term in the natural sciences should not be defined by some people's value judgements.

A third way in which theories of health and disease have been described as value-involving concerns the justification of a theory, rather than the theory itself. As DeVito (2000, 541-542) describes this sort of value-involvement, values enter at the level of choice of criteria, rather than in the criteria themselves. Kingma (2014) explains this type of value-involvement as follows:

> […] the charge is that *even if* a definition of health or function in completely value-free terms succeeds, that definition asserts certain parameters […] that may be *describable* in value-free terms, but cannot be *justified* in a value-free way; they might have been stated differently, and using some rather than others may be driven by, and thus reflect, a normative judgement or evaluative choice. (2014, 600, italics are original)

When discussing this type of value-involvement, Kingma and DeVito target theoretically motivated theories of health and disease that are claimed to be value-free. In Chapter 2 (Section 2.3.1), I will discuss to what extent their charge about value-involvement may actually threaten theoretically motivated theories.

1.2 Desiderata

Having clarified what type of theory that I am interested in developing, let us formulate some desiderata for such a theory. I will present six desiderata, which I will make use of throughout the thesis. Several of these desiderata are related to the discussion in Section 1.1.

From now on, I will use the term 'pathology' instead of 'disease', since this is the term that the intended theory will define.

1.2.1 Theoretical Soundness

A first desideratum is that the theory should be theoretically sound. This desideratum is important to state, because it may come into conflict with making the theory practically applicable. The main reason that I favor theoretical soundness over practical applicability is that I am interested in whether it is possible to define 'health' and 'pathology' in a way that makes sense in medical theory. This is a more basic question than whether these notions may be

defined in a practically applicable way. In order to formulate practically applicable definitions, one should first be clear about whether those definitions actually (at least roughly) track what they are intended to track. However, if there are several alternative definitions of 'health' and 'pathology' that are as theoretically sound, it may be a good idea to opt for the one that is more practically applicable.

1.2.2 Only Empirical, Statistical, and Logical Terms

Since I work under the assumption that health and pathology are natural properties described in the biological disciplines of physiology and pathology, the theory should not use any "non-natural" terms. Rather, it should only use empirical, statistical and logical terms.

1.2.3 No Value-Involvement

As also already indicated (Section 1.1.10), since I aim at definitions of purely theoretical terms, the theory should not itself involve values. With regard to values motivating the interest for the theory's target, the important point is that these values should not be morally dubious. I will say more about this in Chapter 2 (Section 2.3.1) when discussing the above-mentioned argument by DeVito and Kingma.

1.2.4 Clarity

It is of course generally desirable that a theory is clear. There are, however, reasons to highlight this explicitly as a desideratum. First, because of the explicative aim of the theory that I develop, clarity is of extra importance. If it is not clear what the theory says, then the definitions of 'health' and 'pathology' will simply not serve very well as explications. Second, considering theories closely related to the one that I develop (here I am primarily thinking of the theories developed by Boorse (1977) and Hausman (2012)), rather central parts of these theories are left unsettled. Besides leaving it unclear how to apply these theories in specific cases, this also makes it difficult to evaluate them.

1.2.5 Reference Class-Relativity and Integration

In line with my argumentation in Section 1.1.6, the theory should account for health and pathology as reference class-relative properties. This means that an

organ token's health status should be evaluated against the same standard as every other token of that organ type in the reference class. It also means that if there are two exactly alike organ tokens within a reference class, these must have the same health status regardless of which token organisms they are part of.

The theory must also, as discussed in Section 1.1.6, account for the importance of the integration of different physiological functions for health. These two requirements, we saw, may seem to stand in opposition. In order for a theory of health and pathology to be successful they have to be shown to be compatible.

1.2.6 Reasonable Implications

The theory of health and pathology must of course have reasonable implications. As noted in Section 1.1.2, since I aim at explicative definitions of 'health' and 'pathology', the results of the theory need not harmonize perfectly with judgements in physiology and pathology. However, in cases of disharmony between the theory and judgements in physiology and pathology, the results of the theory must be possible to motivate with reference to clarity, consistency, and theoretical fruitfulness. Still, states thought of as paradigmatic examples of health or pathology should be given extra attention in the evaluation of the theory. Let us consider three such types of states on which the debate has focused.

First, we have states of common diseases. Schwartz (2007, 375) provides some examples of common diseases that are all clearly classified as pathological in medical theory. One is a certain dysfunction of the hip joint in dogs (canine hip dysplasia), estimated to be present in 30 percent of the population in some breeds. Another is urinary dysfunction in humans due to benign prostatic hypertrophy, estimated to occur in more than 17 percent of men older than 70. Still another example is senility of the Alzheimer's type, which affects 16 percent of people older than 85. Another example is from Boorse (1977, 566), who mentions dental caries as an example of a common state classified as pathological by pathologists.

The reason that the debate has focused on these sorts of conditions is that Boorse's theory of health and pathology is not able to account for the pathology of states that are common. This is because Boorse draws the line between health and pathology by statistical means. On Boorse's theory, specific types of pathological states are therefore by definition uncommon. Schwartz's examples show that this implication of the BST is not acceptable. Rather, a

sound account of health and pathology should be able to classify states that are common as pathological. I will discuss "the problem of common diseases" in relation to Boorse's theory in detail in Chapter 2 (Section 2.3.4), and then, in Chapter 6 (Section 6.1), further discuss how to solve the problem.

Second, we have states of situation-specific functioning. As pointed out by Kingma (2010), some physiological functions are adequate to perform only in certain situations. For example, it is adequate for the digestive system to digest when there is food in it, but not otherwise. And for many organs and systems, the adequate quantity (e.g. intensity) of the performance of the physiological function differs depending on the situation.[10] For example, although it is always adequate for a heart to beat, what is an adequate beating frequency of the heart differs depending on the situation. It will for instance differ between occasions of resting, walking and running. This means that the theory of health and pathology must be able to judge an organ healthy, even though it does not perform its function (maximally), when being in a situation where it is not adequate for it to perform that function (maximally).

Third, we have states of situation-specific diseases. Many conditions classified as pathological in medical theory are such that they are typical in certain situations. Here, Kingma (2010) provides the following examples: liver failure in situations of overdosing on paracetamol, scurvy in situations of vitamin C deficiency, and fractures in situations of external violence. And, as Kingma (2010, 253) also points out, these examples are not borderline cases of disease. Rather, many examples of situation-specific pathology are considered paradigm cases of pathology. Hence, a theory of the theoretical concepts of health and pathology should classify (at least many of) these states as pathological.

A reason that the debate has focused on situation-specific functioning and situation-specific pathology is that it has been argued that theories of the sort I aim to develop (defining 'health' and 'pathology' as natural properties in a value-free way) cannot account for both. According to Kingma (2010), such theories face a dilemma: they have to choose whether to account for the situation-specificity of physiological functions, or whether to account for situation-specific pathology. Kingma primarily directs her argument against Boorse's theory, which defines 'health' and 'pathology' in terms of statistical typical functioning. In rough terms, Kingma's specific objection is as follows. In order for Boorse's theory to account for the situation-specificity of physiological functions, what counts as a healthy performance of a physiological

[10] Here I follow Kingma (2010, 245) who distinguishes between "qualitative function" and "quantitative function".

function must be situation-relative. So, a healthy performance of a physiological function in a certain situation is determined by what is a typical performance of that function *in that situation*. However, Kingma argues, this means that it will not be possible for Boorse's theory to account for pathological states that are typical in certain situations. Even if the concerned state is overall untypical, it is typical given the specific situation – and hence healthy. I will discuss Kingma's argument against Boorse's theory in more detail in Chapter 2 (Section 2.3.3), and then, in Chapter 6 (Section 6.2), further discuss how to solve it.

1.3 Outline of the Thesis

The project of defining 'health' and 'pathology' as purely theoretical terms in physiology and pathology will be structured as follows. In Chapter 2, I discuss the most debated theoretically motivated theory of health and pathology – Boorse's "biostatistical theory" – and argue that it does not meet all of the desiderata listed above. Then, in Chapter 3-5, I develop an alternative theory of health and pathology, which succeeds better than Boorse's in satisfying the desiderata. I call this theory 'the disposition profile efficiency theory'. The presentation of the disposition profile efficiency theory is arranged in the following way. Chapter 3 introduces the theory by laying out some basic assumptions and defining what may be described as the theory's core part – "disposition profiles". Chapter 4 is connected to the desideratum that health and pathology should be accounted for as reference class-relative properties. It defines 'reference classes' and accounts for reference class-relative standards for health. In Chapter 5, I provide a measure of efficiency, and use this to define 'health' and 'pathology'. Chapters 6 and 7 discuss implications of the disposition profile efficiency theory. Chapter 6 concerns the problem of common diseases and Kingma's dilemma, which I referred to in the above section about reasonable implications. Chapter 7 discusses contested types of conditions mentioned in Section 1.1.3, namely conditions that are trivial from the perspective of the whole organism, normal ageing, defense mechanisms and risk conditions. Chapter 8 concludes. There is also an appendix which collects all definitions of the disposition profile efficiency theory. This collection may be useful since the development of this theory spans over three chapters.

For the presentation of the disposition profile efficiency theory, I will make use of formal notation. Some central parts of the theory will also be formally modelled. The motivation for this is, first, clarity. It should be made precise

what the theory says. The motivation is, second, concision. Without the use of formal notation and formal models, complex definitions become rather cumbersome and hence more difficult to grasp.

2 The Biostatistical Theory

In this chapter I will discuss the biostatistical theory (abbreviated 'the BST'), developed and defended by Boorse in a number of papers (1976b; 1977; 1987; 1997; 2002; 2011; 2014). The biostatistical theory is probably the most debated theory of health and disease, or pathology. I will start by describing Boorse's aim (Section 2.1), and then continue by discussing the content of the theory (Section 2.2). Last, I will discuss four types of objections posed towards the biostatistical theory, which will be important for the development of my alternative theory of health and pathology in the next chapters (Section 2.3).

2.1 Aim

As we saw in Chapter 1, the most basic question for a theory of health and disease is what motivates it. Boorse's motivation for the biostatistical theory is purely theoretical. The BST is about the concepts of health and disease in medical theory, which Boorse understands as physiology and pathology. Boorse distinguishes between these theoretical, value-free concepts of health and disease and other, normative concepts of health and disease, for example clinical, political and social concepts, which the BST is not about (1997, 45-46). Given this aim, facts about how 'disease' and 'health' are used in everyday language are largely uninformative. Instead, physiologists and pathologists' usage of the concepts plays an important role in evaluating the account. Boorse writes:

> Our main task now is to analyze the normal-pathological distinction in traditional medicine. A correct analysis, or reportive definition, must conform to medical usage–that is, it must fit the stock of recognized pathological conditions. Definitions that are wider or narrower than this stock are incorrect. (1987, 366)

In the above quote, Boorse seems to aim at descriptive definitions of 'health' and 'pathology'. Other passages, though, suggest otherwise. Boorse (2014, 713) indicates that his project is not just descriptive, but possibly explicative. On this theme, he also writes:

> I am content for the BST to live or die by the considered usage of pathologists– which does not, of course, exclude that on reflection (as in Rawlsian equilibrium), pathologists might revise their usage slightly to achieve consistency with a simple and powerful theory. (Boorse 1997, 53)

In his latest defense of the biostatistical theory, Boorse (2014) more explicitly comments on what definitions the BST aims at:

> […] Nordenfelt is right (2001, 26) to describe my project as "philosophical explication," or rational reconstruction, in the logical-empiricist tradition of Hempel, Carnap, and Quine. […] I still think my effort is in some sense an attempt at a lexical definition of scientific terms (Boorse, 1977, 551), but with the provisio that scientists are sometimes confused, inconsistent, or (as with fever) empirically wrong about their subject. (2014, 713)

These three quotes illustrate somewhat of an oscillation in Boorse's writings between a more descriptive project on the one hand and a more explicative project on the other.[11]

The BST is, as stated above, an attempt to define the theoretical terms 'health' and 'disease'. To what do these terms refer? Although, as we saw in Chapter 1, the exact extension is a matter of controversy, we can at least ask what rough idea of the extension Boorse has in mind. Boorse uses the claim, or "traditional axiom of medicine" as he calls it, that "health is the absence of disease" as a presupposition when developing the BST (1977, 542). Like most other theorists, Boorse takes disease not only to include conditions that we in everyday language call 'diseases', but all conditions that are not healthy, for example injuries, poisonings, and growth disorders (1987, 362-363; 2011, 26). However, Boorse only claims the BST to be about somatic conditions, not mental conditions. But he does not rule out that his definitions of 'health' and 'disease' apply to mental conditions; rather, he leaves that an open question (Boorse 2011, 29).[12] Also, it should be pointed out that Boorse thinks of health and disease as universal biological phenomena. The BST is supposed to cover

[11] See also Schwartz (2014) on this issue.

[12] Boorse discusses mental health in a separate paper (1976a). The BST is, however, independent of that discussion.

health and disease not only among humans, but all organisms. Boorse writes that "[w]hat a healthy hen or cow is like is a biological fact; it is not an economic one" (1977, 565). In relation to this, we may also note that Boorse thinks of health and disease as natural properties. Thereby, he also accounts for health and disease in empirical and statistical terms only.

In his first papers about the BST, Boorse defines 'health' and 'disease'. Later, Boorse switches from 'disease' to 'pathology' or 'pathological condition'. This is, however, only meant as a terminological change. Boorse thinks that 'disease' might be misleading with regard to what he aims to account for, and that 'pathology' fits his purposes better (1997, 7; 2014, 684). Two related reasons for this are (i) that 'pathology', in contrast to 'disease', is a term used by scientists rather than lay people, and (ii) that 'pathology' is taken to have a wider extension than 'disease' (including all conditions that are not healthy). Boorse also switches from 'health' to 'normal' (1997, 7). But unlike the switch to 'pathology', Boorse does not consistently switch to 'normal'. Rather, the discussion is often conducted in terms of the distinction between 'health' and 'pathology'. In the following, I will use 'disease' and 'pathology' synonymously. I will, however, stick to the term 'health'. This is in order to avoid confusion: Boorse uses the adjective 'normal' in several contexts ("normal function", "normal efficiency level", "normal environment") where the term does not refer to the same thing as 'healthy', but rather to something like "statistically typical".

Something Boorse stresses about the BST is that it is a value-free theory (1977, 542-543). The concepts analyzed, he says, belong "to the life-death family of biological concepts, not to the welfare-harm family of evaluative ones" (Boorse 2011, 28). The theory has no normative components and no normative implications by itself. Hence, it cannot be used (in isolation) to answer practical questions, for instance whether a certain physiological state is good or bad, or whether it should be treated.

2.2 Content

The biostatistical theory is built on four major components: (i) reference classes, (ii) physiological functions, (iii) efficiency, and (iv) statistical normality. Boorse uses these components to define 'health' and 'pathology' roughly as follows: An organ in an organism is healthy if and only if the organ is able to perform its *physiological functions* with an *efficiency* that is at least *statisti-*

cally normal for organs of that type in the organism's *reference class*. Otherwise the organ is pathological. Beneath, I will explore the idea of the BST in detail by discussing each of the four components in turn.

Usually, the BST is described as consisting of only three components, leaving out efficiency. This component is instead commonly treated as part of the component of statistical normality or as part of the component of physiological functions. However, as we will see below, there are issues regarding both how Boorse's ideas of efficiency, physiological functions, and statistical normality are to be interpreted by themselves, and how they are to be interpreted in relation to each other. Therefore, I think it is useful to discuss efficiency as a separate component.

2.2.1 Reference Classes

A basic idea of the biostatistical theory is that health and pathology are relative to reference classes. The health status of an organism is determined by comparing the organism to a statistical ideal for that organism's reference class. For example, when determining the health status of a newborn human male baby, we have to consider this individual in relation to what is typical for its reference class – which, on the BST, consists of human male babies.

According to Boorse, reference classes are "natural class[es] of organisms of uniform functional design" (1977, 562). Boorse motivates this general idea of reference classes, first, by observing comparative physiology. He claims that comparative physiology describes ideal types of organisms, for example "the frog, the hydra, the earthworm, the starfish…" (Boorse 1977, 557). Boorse calls these ideals 'species designs'. He takes them to be statistical abstractions of the members of the species. "Each detail of this composite portrait [i.e. species design] is statistically normal within the species" (Boorse 1977, 557). Boorse, however, points out that it might well be the case that no member of the species fully instantiates the species design, and rather that everyone is statistically atypical in at least some respect (1977, 557).

Second, Boorse motivates the idea of reference classes by referring to evolutionary biology. Although evolutionary biology emphasizes constant variation, Boorse points out that "[t]he typical result of evolution is precisely a trait's becoming established in a species …" (1977, 557). "Normalizing selection", Boorse argues, causes species designs to arise. Although these designs are altered through evolutionary history, there is a "short-term constancy", and it is on this, Boorse thinks, that medical theory and practice rely (1977, 557).

Third, Boorse motivates the idea of reference classes with the observation that, in medical practice, assumptions are constantly made about patients being very similar to each other. For example, when diagnosing and treating a patient with pancreatitis, it is taken for granted that the patient, just like other patients, has a pancreas located near the stomach, and that an inflammation of this individual's pancreas would cause certain symptoms, just like in other patients with pancreatitis (Boorse 1977, 557).

When accounting for how reference classes are individuated, Boorse refers to physiological theory. First, in accordance with comparative physiology, Boorse takes reference classes to be individuated by species. Second, considering descriptions of species, Boorse thinks that reference classes should be individuated by sex and age. This is because the sets of physiological functions for females and males differ (think for example of ovulation and sperm production). So do the sets of physiological functions for organisms of different ages or life-stages (think for example of enlargement of the skeleton, or again about ovulation or sperm production) (Boorse 1977, 558). Therefore, Boorse takes reference classes to be relative to species, sex, and age. He writes that "[i]n medical applications the operative class seems to be an age group of a sex of a species, e.g., human male neonates or, say, 7-9 year old girls" (Boorse 1977, 558). Boorse also mentions that "[i]n other contexts [than medical applications], perhaps even in medicine itself, one would have to factor in race as well, since in some respects the different races have different functional designs" (1977, 558). Although reference classes are more fine-grained than species, Boorse uses the term 'species design' to denote ideal designs for reference classes, since he thinks that "it is still convenient and unlikely to cause confusion" (1977, 558).

The members of a reference class need, however, not share a uniform functional design in a strict sense: Boorse allows for some functional variation within the design of the reference class. With regard to some types of organism parts, there are so called polymorphisms, i.e. rival variants that are all typical in a population. As examples of polymorphisms Boorse mentions the four human blood groups (A, B, AB, and O), the three different colors of the human iris (brown, blue, and green), and the different amount of pigmentation in humans (from small to large). Boorse includes such functional variants disjunctively in the species design (1977, 558). Hence, on Boorse account, the blood group of the reference class design for 7-9-year-old human females is (A or B or AB or O).

A further question about reference classes regards their extension in time. According to Boorse (1997, 66), reference classes do not exclusively contain

organisms alive at a certain time. Rather, reference classes comprise organisms from a "time-slice of a species" (Boorse 2014, 715). We may note that this view on reference classes fits well with Boorse's comments about reference classes and evolution above, where he claims that although species designs are altered through evolutionary history, there is a short-term constancy on which medical theory and practice rely.

2.2.2 Physiological Functions

As we saw above, Boorse describes members of a reference classes as having a "uniform functional design". This means that they are alike in their physiological functions. Let us now consider physiological functions: What kind of entities have physiological functions? And what does it mean to say of such an entity that it has a physiological function?

Boorse uses different terms to denote entities that have physiological functions. One is the biological term 'trait', which denotes a characteristic or feature of an organism. For example, Boorse says that "[b]iologists regularly use functional language to describe the role of traits in the life of organisms" (1977, 555). But mostly Boorse uses other terms. He talks about functions of "internal parts" of organisms. He also uses the terms 'parts' and 'processes' (see e.g. Boorse 1977, 555, 558). As concrete examples of parts and processes, Boorse mentions the heart and the kidney (1977, 556). He also mentions smaller entities, for example organelles and cells (Boorse 1997, 7). In the remainder of the presentation of the biostatistical theory, I will for reasons of generality and a uniform language use, when possible, talk about physiological functions of traits.

How traits are individuated is something that Boorse does not discuss. But he allows for traits at different levels. For example, if we look at the circulatory system as a trait, we find the heart as a lower-level trait, and the mitral valve as an even lower-level trait.

Let us then turn to the question what it means for a trait to have a physiological function. The view of functions that Boorse defends unifies function statements about artifacts and function statements about biological entities. What makes artifacts as well as biological entities have functions, according to Boorse, is that they causally contribute to some goal of a system they can be described as parts of. More precisely, the analysis goes:

> *X* performs the function *Z* in the *G*-ing of *S* at *t* if and only if at *t*, the *Z*-ing of *X* is a causal contribution to *G*. (Boorse 2002, 70)

Here, *G* is a goal, *S* a system, and *t* an interval of time. Hence, according to Boorse, function statements are relative to: first, a system of which the thing performing the function is part; second, a goal of this system; and, third, a time of performance. We may also note about this analysis that it makes function statements relative to a decomposition of the system (where *X* is a part).

According to Boorse, it is especially in the determination of the goal that functions of artifacts and functions of biological entities differ. The goals that artifacts contribute to are subjectively determined in the sense that they are set by human (or other) minds (Boorse 2002, 68). For example, an eraser (usually) has the function to rub out marks of pencils, since that is the (usual) intention of an eraser user. But, also, if someone puts an eraser on a certain line in a book to remember where to start reading the next time, then that eraser has the function of serving as a bookmark.

The goals that biological entities contribute to are, in contrast to the goals to which artifacts contribute, objectively determined, Boorse argues (1976b, 79). The goals are objectively determined, Boorse thinks, because biological entities are inherently goal-directed. They are so as a result of evolution, which "seems to yield organisms with the supreme goals of individual survival and reproduction" (Boorse 2002, 64). Boorse expresses the idea of goal-directedness in the following way:

> To say that an action or process *A* is directed to the goal *G* is to say not only that *A* is what is required for *G*, but also that within some range of environmental variation *A would have been modified* in whatever way was required for *G*. (1976b, 78, italics are original)

Importantly, though, Boorse points out, biological entities do not strive towards goals in any intentional sense:

> Biologists tell us that 'plant exhibit heliotropism, that is, turn in varying directions *in order* to maximize their exposure to the sun. But no biologist is prepared to find the causes for this behavior in a plant's belief about the role of photons in photosynthesis and in its desire to maximize the number of photons landing on its leaves' (Rosenberg 1985: 44). Thus, we expect the purposiveness common to living organisms' behavior to be analyzable in some non-mental, but naturalistic, way. (2002, 68-69, italics are original)

Boorse's idea that organisms are goal-directed is grounded in his teleological view of the biological realm. Boorse mentions that teleological explanations

in the area of biology are frequently viewed with suspicion. These misgivings are, however, often excessive, he thinks (Boorse 1977, 554).

Another characteristic of biological entities that Boorse emphasizes, which relates to their goal-directedness, is that they typically exhibit an extensive means-end hierarchical organization (2002, 69-70). The biological realm comprises a lot of intertwined systems at different levels with different goals:

> Individual cells are goal-directed to manufacturing certain compounds; by doing so they contribute to higher-level goals like muscle contraction; these goals contribute to overt behavior like web-spinning, nest-building, or prey-catching; overt behavior contributes to such goals as individual and species survival and reproduction. (Boorse 1977, 556)

So, depending on what biological system in the hierarchy we look at we will locate different goals. Since various biological disciplines (e.g. genetics, physiology, ecology) have different systems as their interest, Boorse thinks that they also consider different goals. Although there may be many levels of subsystems with their own local goals, the function statements made in a biological discipline, Boorse explains, will ultimately be relative to the highest-level goal(s), i.e. the goal(s) of the goal-directed system of interest. In physiology, Boorse claims that the system of interest is the individual organism. And the goals of the individual organism, he claims, are its own survival and reproduction (Boorse 1976b, 84). Importantly, though, for a function to count as a physiological function, it only needs to contribute to one of these two goals (Boorse 2014, 685). Hence, according to Boorse, physiological functions are functions that contribute either to the organism's individual survival, or to its individual reproduction, or to both.

How should we understand "individual survival and reproduction" here? Boorse does not give any precise definition of these goals. But, at least, Boorse seems to think that these goals are rather similar to the concept of fitness in evolutionary biology. In a general discussion about goal-analyses of biological functions, Boorse writes:

> Since evolution in fact seems to yield organisms with the supreme goals of individual survival and reproduction (loosely, 'fitness'), within biology the GGC analysis [general goal contribution analysis] gives the same result as one defining biological functions, specifically, as causal contributions to fitness. (2002, 64)

Still it is not clear exactly how to understand the goals of individual survival and reproduction. Especially the goal of individual reproduction merits discussion. The term 'individual reproduction' gives the impression of only including offspring begotten by the individual itself. It hence seems to stand in contrast to the concept of inclusive fitness, which also includes offspring begotten by other individuals that have genes in common with the individual in question. Boorse also writes "[…] a trait which contributes only to other goals like the survival of the species–for example, the immolation response of some ants to fire–will be omitted from physiological theory" (1976b, 84). However, in later writings, Boorse expresses himself positively towards counting inclusive fitness as a goal. In a footnote, he writes:

> […] kin selection […] could be viewed as a type of reproduction and thus as normal by the BST. Although I know no medical basis for such a judgement, it seems a fairly natural extension of medicine's existing views on reproduction.[13] (Boorse 1997, 124)

And when discussing what the BST implies about homosexuality, Boorse writes:

> But if, for example, any of the kin-selection hypotheses were correct, by the BST homosexuals would be as normal as worker bees. They would merely have a variant method of reproduction. I fully accept inclusive fitness (Garson and Piccinini, 2014) as a biological goal. (2014, 691)

Leaving the issue of how to understand individual survival and reproduction aside, there is more to say about the role the goals play in determinations of physiological functions. Although physiological functions are functions that contribute to organisms' individual survival or reproduction, not all functions performed in an organism that contribute to its survival or reproduction are physiological functions, on Boorse's view. Boorse says: "[w]hatever contributes to these goals [individual survival and reproduction] *reliably, throughout a species or other reference class*, is assigned a physiological function" (1976b, 84, my emphasis).[14] In this quote, we can see that physiological functions have two related features, which I will explain below: (i) they are type-

[13] I think we should understand "kin selection" in this citation as reproduction by close relatives.
[14] The reason Boorse writes "a species or other reference class" rather than merely "reference class", I guess, is that many physiological functions are shared by all reference classes of a species.

relative (rather than token-trelative), and (ii) statements of physiological functions are strong function statements (rather than weak function statements).

Let us first look at the type-relativity of physiological functions. Think again about an eraser, which during a certain time interval is used to rub out marks of pencils on a paper. By contributing to the goal of cleaning the piece of paper, at the time in question, this token eraser has the function of rubbing out marks of pencils. The kind of function here, i.e. a token-relative function, is what was given an analysis above. Observe that a token may be only a time-slice of a persisting object. For example, since a certain eraser might not always be used to rub out marks of pencils, we may have to view the persisting eraser as a series of several token erasers existing at different times (some of them having the function of rubbing out marks of pencils).

In contrast, we may also make a claim about the eraser, as a type of entity, that it has the function of rubbing out marks of pencils. Physiological functions, as Boorse describes them, are of this type-relative kind. When physiologists claim that the heart has the function of pumping blood, this is not primarily a claim about a single heart token, but about the heart as a type of organ.

In virtue of what does a type have a function? We get an answer to this by looking at Boorse's distinction between "weak" and "strong" function statements. By weak and strong function statements, Boorse takes a view in the discussion about whether there is a substantial distinction to be made between "accidental functions" and "proper functions". Especially in the biological realm there are tendencies to think of some functions as proper functions (or "functions possessed", or "real functions"). For example, the function of pumping blood is usually thought of as a proper function of the heart. In contrast, a heart's function of making noises, in a situation where a physician is examining the heart with a stethoscope, is usually not thought of as a proper function of the heart.

Whereas philosophers like Millikan (1984) and Neander (1991) argue that proper functions are of a special kind, not to be confused with accidental functions, Boorse (1976b, 81) argues that there is no such strong distinction to be made. However, there is a difference in how often the function is performed. What makes us conceive of some functions as proper rather than accidental, Boorse (2002, 71) explains, is that they are more frequently performed: The function of pumping blood is usually thought of as a proper function of the heart, since heart tokens frequently contribute to their bearers' individual survival by pumping blood. In contrast, the function of making noises is usually not thought of as a proper function of the heart, since making noises is not an

activity by which heart tokens frequently contribute to their bearers' individual survival.

In Boorse's terminology, this difference means that the first function statement is a strong function statement and the second function statement is a weak function statement. Whereas weak function statements are usually expressed by the terms "*X* performs the function *Z*" (as in the analysis given above), strong function statements are usually expressed by the terms "the function of *X* is *Z*", or "a function of *X* is *Z*" or "*X* has the function *Z*" (Boorse 2002, 71).

Physiological function statements, according to Boorse, are strong type-relative function statements. As we saw, Boorse holds that physiological functions are functions that contribute to individual survival or reproduction reliably, throughout a reference class. This is taken by Boorse to imply, for example, that the heart has the physiological function of pumping blood, but not of making noises. However, if it became common for heart tokens to contribute to their bearers' survival by making noises, making noises would become a physiological function of the heart.

In line with the above presentation, Boorse talks about the distinction between weak and strong function statements only as a difference in how often, or regularly, a function is performed among the tokens of a type. However, when responding to an objection to his analysis of physiological function by Millikan (1993), Boorse (2002, 92-93) seems to have a more complex idea in mind of how physiological functions are determined. According to Millikan's objection, Boorse's goal-analysis excludes some physiological functions – those that are performed only by very few tokens of a type. An example by Millikan are sperm tails, which have the function of propelling the sperm to an ovum, although most sperm tails never do so. In his answer to Millikan's objection, Boorse refers to "specific occasions" for performances of functions:

> The key point is that most biological functions are performed only on specific occasions. Failure of a sperm to fertilize a non-existent ovum is like failure of blood to clot in a non-existing wound, or failure of sweat glands to release sweat when core body temperature is not above a certain level. A great many functions are occasion-specific; the occasions may be rare. (2002, 93)

The implicit idea here seems to be that physiological function statements are strong function statements in the sense that the function is frequently performed by the tokens of a type given specific circumstances. Sperm tails have the function of propelling the sperm to an ovum, according to Boorse, since

they frequently do so in the right circumstances, i.e. when there is an ovum nearby in the right direction, or something along those lines.

Boorse is less than completely clear and consistent when discussing his analysis of physiological function. But, if we try to take into account what he says in his answer to Millikan, and try to make the theoretically most sound definition of 'physiological function' out of it, I suggest the following:

A function F is a physiological function of a trait type T if and only if:

(i) there is some set of circumstances CIR, in which F is performed by many tokens of T, and
(ii) in CIR, the performance of F by a T-token contributes to the survival or reproduction of the T-token's bearer.[15]

A last question about physiological functions is how they are specified. When talking about physiological functions, Boorse points out that what he has in mind is not "the concrete process that makes a physiological contribution". Rather, what he has in mind is "the contribution to physiological goals". Taking the thyroid as an example (which by secreting thyroid hormones regulates the organism's overall energy consumption), Boorse writes:

> In one sense, sometimes used with clinical tests, a function is the concrete process that makes a physiological contribution, e.g. thyroid secretion. […] This is not our usage, since for us the function is the contribution to physiological goals, and too much thyroid secretion damages these goals as much as too little. To put it another way, the function of the thyroid is not merely to secrete hormones, but to secrete the right amount of them for current metabolic needs. (1977, 559)

But how narrowly should we then understand physiological functions? For example, consider mechanoreceptors. Mechanoreceptors are a type of sensory receptors which are sensitive to mechanical forces. When there is a mechanical pressure or distortion these receptors recognize this and send a signal to the central nervous system (CNS), through afferent neurons. What should we say is the physiological function of the mechanoreceptors? Is it to recognize and send signals about pressure and distortion to the CNS, or is it rather to

[15] This is similar to the account of physiological function by Garson and Piccinini (2014), intended as an improvement of the BST.

recognize and send signals about pressure and distortion with a certain (adequate) intensity to the CNS, or rather to recognize and, without too much delay, send signals about pressure and distortion with a certain (adequate) intensity to the CNS? When Boorse says that "the function of the thyroid is not merely to secrete hormones, but to secrete the right amount of them for current metabolic needs" one may interpret him as meaning that physiological functions have very narrow descriptions. However, when we next turn to discuss efficiency, we will see that Boorse wants to be able to compare different performances of the same physiological function with regard to their efficiency (roughly, how "well" the function is performed). This means that Boorse cannot reasonably have too narrow (and "perfect") descriptions of physiological functions in mind. Very narrow descriptions would imply that a thyroid that secretes thyroid hormones close to, but not exactly "the right amount of them for current metabolic needs", does not perform the physiological function of thyroids.

A way to interpret the BST, then, is to specify physiological functions merely by quality and not by quantity. This is a distinction Kingma (2010) uses in order to understand the BST in the most reasonable way. Kingma writes:

> Though the BST states that health is a normal function, merely performing a normal or species-typical function is not enough for health. In order to equate to health, normal or species-typical function must be performed *at the right level*: the heart, for example, must not merely pump blood, but it must pump blood at the right speed, pressure, etc. I call the former function ascription–pumping blood–the *qualitative* normal function of the heart, and the latter function ascription–pumping blood at the right level, speed, pressure, etc.–the *quantitative* normal function of the heart. *Quantitative* normal function is normal or species-typical function at *the right efficiency level* […] (2010, 245, italics are original)

So, although Boorse sometimes seem to indicate otherwise, the most reasonable interpretation of the BST is that physiological functions are not quantitatively specified at all, but only qualitatively.

2.2.3 Efficiency

Considering physiological functions in a reference class, there can be better or worse performances of them with regard to the goals of survival and reproduction. For example, some heart tokens pump blood in a way that supports

survival and reproduction well, whereas other heart tokens pump blood in a way that does not support survival and reproduction as well. In Boorse's terminology, different performances are of different efficiencies. Efficiency, Boorse clarifies, is not to be understood simply as the amount of a certain process carried out. To see this, consider again the thyroid. A very extensive secretion of thyroid hormones does, just like a very limited secretion, reduce rather than enhance survival. Efficiency, Boorse explains, is about the process' serving the physiological goals (1977, 559). Whereas there can be too much of the concrete process, there cannot be a too high efficiency (Boorse 1977, 559). A thyroid performing very efficiently, in Boorse's terminology, is a thyroid that serves the goal of survival very well, i.e. a thyroid that secretes an appropriate amount of thyroid hormones.

It is, however, not clear from Boorse's writings how to understand "efficiency" more precisely. First, Boorse does not suggest any units of measurement for the physiological goals. Units of measurement for the goal of survival could, for example, be chances of surviving for a certain period of time, risk of dying at a certain age or within a certain period of time, expected length of life, or further expected length of life. Second, given such units, it is not clear how to measure how efficiently a trait token's performance serves the goals.

One thing that contributes to making this unclear is that efficiency is taken to regard two different goals at the same time – survival and reproduction. Even if we define separate measures for survival and reproduction, it is unclear how these measures relate to each other within the efficiency measure. The reason for this is that a trait token's performance of a physiological function may serve the possessor's survival better in comparison to the reference class than it serves the possessor's reproduction in comparison to the reference class (or vice versa). For example, think of a human female at age 23, whose pituitary's secretion of hormones is abnormal in a way that makes her ovulate very infrequently. The abnormal hormone levels do, however, not affect her notably in any other way. Here, the pituitary's abnormal secretion of hormones makes her worse off with regard to reproduction, relative to her reference class, but not worse off with regard to survival, relative to her reference class. If Boorse's notion of efficiency is to make sense, it must include some weighing of survival and reproduction. But it is not obvious that there is a reasonable weighing to formulate here.

Since Boorse leaves the notion of efficiency in the biostatistical theory largely unsettled, it is also unclear to what extent the BST accounts for health and pathology as reference class-relative properties. When determining the efficiency of a trait token's performance, should we consider how well the

physiological function is performed given the construction of the particular organism with the trait token under evaluation, or should we rather consider how well the physiological function is performed given some standard organism-construction?

2.2.4 Statistical Normality

It is by comparing the efficiency levels of the performances by different tokens of a trait type that the biostatistical theory draws the line between health and pathology. The thought is that, for each trait type in a reference class, there are normal efficiency levels for the performances of the trait type's physiological functions. "Normal" here means statistically typical. For example, in a reference class of human 7-9-year-old females there is supposed to be a (range of) statistically typical efficiency level(s) for these individuals' hearts' blood pumping. On the BST, healthy trait tokens are associated with performances of at least typical efficiency and pathological trait tokens are associated with performances of less than typical efficiency (Boorse 1977, 558-559, 562). Exactly where along the range of efficiency levels to locate the line distinguishing at least typical efficiencies from less than typical efficiencies is somewhat arbitrary (Boorse 1987, 371), or conventional (Boorse 1977, 559). However, Boorse (1977, 559; 1997, 8) says that the distinguishing line must be drawn a certain distance below the average efficiency, and Boorse (2014, 684) says that it must be drawn far below the average efficiency.

We may note here, in relation to the issue discussed in Chapter 1 of whether the concepts of health and disease are absolute or comparative, that the BST draws an absolute line between health and pathology. Either the efficiency is high enough for the trait token to count as healthy, or it is low enough for the trait token to count as pathological. However, the BST's distinguishing line should not be regarded as too strict; Boorse says that "the concept of a pathological state has vague boundaries" (1987, 371).

It is not clear exactly how Boorse thinks the line distinguishing health from pathology is drawn. But, as I understand him, it is at least "statistically determined". I take "statistically determined" to require that there is a function

$$d : \text{distributions of efficiency levels} \rightarrow \text{efficiency levels}$$

such that, for each distribution E of efficiencies of trait token performances of a physiological function in a reference class, $d(E)$ is the efficiency level where

healthy trait tokens are distinguished from pathological trait tokens. Importantly, one and the same function d is used to draw the distinguishing line in each case. The existence of such a function d is probably not a sufficient condition for deeming the distinguishing line statistically determined. I do not have any concrete suggestions for sufficient criteria, but I think the idea is that d should be a "non-exotic" function defined in terms of statistical measures such as averages and standard deviations. In particular, a definition of d on a case-by-case basis should not qualify as statistical.

One may, perhaps, also conceive of a weaker interpretation of the BST, where the theory does not fully account for how health is distinguished from pathology. Here, the thought would be that Boorse merely provides some necessary conditions that such an account must respect, for example that the distinguishing line must end up below, or far below, the average efficiency level. In particular, this interpretation would allow for non-statistical ways of drawing the line.

I said above that, on the BST, healthy trait tokens are associated with performances with at least typical efficiencies and pathological trait tokens are associated with performances with less than typical efficiencies. In order to state more precisely the BST's definitions of health and pathology we need to complicate things a bit. Typically, trait tokens do not just constantly perform a certain function. Rather, they perform their functions on "appropriate", or "typical", occasions (Boorse 1977, 562). Boorse writes:

> […] vision occurs when the eyes are open, digestion when food is in the alimentary canal, adrenalin secretion under stress, sweating when temperature is rising, blood-clotting after a wound, and so on. (1977, 562)

What is healthy is not simply to perform a certain function (by a certain quantity). Rather, it is to have the "readiness" to perform a certain function (by an adequate quantity) on the right occasions. In order to take this into account, Boorse defines 'health' in terms of *dispositions* of trait tokens to perform their functions *on typical occasions*. This means that an organism that is not in a situation that is typical for the performance of a certain function does not count as having a pathological condition merely because it does not perform that function. For example, the optic nerve does not become pathological as soon as the bearer closes her eyes. It also means that an organism that does not have the ability to perform a certain physiological function is not considered healthy just because she avoids situations that are adequate for performances

of that physiological function. For example, "[h]emophiliacs who are protected from all injury, or diabetics who take daily insulin, are still diseased" (Boorse 1977, 562).

By thinking of health and pathology in this dispositional way, we get the following definitions:

> A trait token is **healthy** if and only if, for each of its trait type's physiological functions, it is disposed to perform the function on typical occasions with at least typical efficiency for the trait token's bearer's reference class.

> A trait token is **pathological** if and only if, for at least one of its trait type's physiological functions, it is not disposed to perform the function on typical occasions with at least typical efficiency for the trait token's bearer's reference class.

Or, if we look at Boorse's own formulation of the biostatistical theory:

> 1. The *reference class* is a natural class of organisms of uniform functional design; specifically, an age group of a sex of a species.
> 2. A *normal function* of a part or process within members of the reference class is a statistically typical contribution by it to their individual survival [or] reproduction.
> 3. *Health* in a member of the reference class is *normal functional ability*: the readiness of each internal part to perform all its normal functions on typical occasions with at least typical efficiency.[16]
> 4. A *disease* [later, *pathological condition*] is a type of internal state which impairs health, *i.e.,* reduces one or more functional abilities below typical efficiency. (2014, 684, italics and brackets are original)

We may note here that Boorse defines pathology of trait tokens (or internal parts), but health of whole organisms. By defining pathology of trait tokens, Boorse clearly takes a compositionalist view on pathology. What about health, then? Although he defines the health of a whole organism in the clauses above, his view on health may also be described as compositional rather than holistic.

[16] As is clear in Boorse's writings (and as I have described the BST above), the idea is that an organism is healthy if each of its trait tokens has the disposition to carry out *its respective type's* physiological functions with typical efficiency. So, the third clause would be more adequately expressed in the following way: Health in a member of the reference class is normal functional ability: the readiness of each internal part to perform all of its type's normal functions on typical occasions with at least typical efficiency.

This is because the health of an organism is not defined by any feature of the organism as a whole, but of the organism's separate trait tokens: a member of a reference class is healthy if each of its trait tokens is not pathological – i.e. healthy.[17]

When discussing Boorse's notion of efficiency, I noted that it is unclear whether it should be understood as relative to the individual organism under evaluation, or relative to some common standard for the reference class. By drawing the line between health and pathology statistically in the reference class distribution of efficiency levels, it may seem that the BST, regardless of how we understand efficiency, accounts for health and pathology as reference class-relative properties. This is, however, not the case. One may say that, because of the BST's distinguishing line, it accounts for health and pathology as reference class-relative properties in a weak sense: which efficiencies indicate health and which efficiencies indicate pathology depend on the distribution of efficiencies in *the reference class*. However, for the BST to account for health and pathology as reference class-relative properties in the strong sense described in Chapter 1 (Section 1.1.6), it is crucial that the notion of efficiency is relative to reference classes. That the notion of efficiency is reference class-relative means that the efficiency of a trait token performance is about how well the physiological function is performed given some standard organism-construction, and not about how well the physiological function is performed given the construction of the particular organism. Without a reference class-relative notion of efficiency, two functionally identical heart tokens may have different efficiencies and hence be differently located in the reference class distribution of efficiencies. This would mean that it is not guaranteed that two identical trait tokens of the same type within the same reference class have the same health status.

2.3 Objections

I have now presented the biostatistical theory in its entirety and discussed several unclear parts. Beneath, I will consider some further objections towards the theory. I will first consider objections towards the BST's individuation of reference classes (Section 2.3.1) and Boorse's choice of survival and reproduction as the relevant physiological goals (Section 2.3.2). I will then discuss

[17] To be more accurate, we may note that there are also holistic aspects built into Boorse's definitions of 'health' and 'pathology' – namely the reference to the physiological goals of the organism. See Täljedal (2004) for further discussion on this topic.

"Kingma's dilemma" (Section 2.3.3), and "the problem of common diseases" (Section 2.3.4). My reason for discussing these objections are that they (i) are topics in the current debate, (ii) have not been sufficiently and satisfyingly answered, and (iii) are central for the theory that I develop in the next chapters. Lastly, I will note that there are additional objections to the BST, which I will not discuss in this thesis (Section 2.3.5).

2.3.1 Individuation of Reference Classes

A component of the biostatistical theory which has been much discussed is that of reference classes. Remember that Boorse accounts for reference classes as natural classes of organisms, individuated by species, sex, and age. Several authors observe that depending on how reference classes are individuated, different conditions count as pathological. For example, DeVito writes:

> There are many other possible reference classes one could choose and, depending on what one chooses as a reference class, different body states will be classified as healthy and diseased. For example, imagine that we have a population of 70-71-year-old women all of whom have osteoporosis. When these women fall they invariably break their hips. Are their bones functioning abnormally or normally? If we take Boorse's choice of reference class (species, age, sex), then they are all properly functioning. But if we instead remove (or broaden) the age restriction on the reference class, we get a different result. Looking at women between the ages of 13 and 85, we find that osteoporotic bones are not statistically normal. In addition, if we take the individual 70-year-old's life as the reference class, then osteoporotic bones are also not statistically (or historically) normal. Thus health and pathology are dependent upon how we group individuals. (2000, 546)

Further illustrations of how decisive the individuation of reference classes is for the classification of states as healthy or pathological are made by Kingma (2007; 2014). She points out that one gets different extensions of 'pathology' if one individuates reference classes only by species, sex, and age, compared to if one in addition individuates reference classes by sexual orientation, certain chromosomal variants (e.g. having an "extra" chromosome 21), or alcohol consumption. By the first definition, homosexuality, Down's syndrome and liver cirrhosis come out as pathologies (in all human reference classes) by reducing the chances of survival or reproduction of the concerned individual in comparison to the reference classes. By the alternative definitions, on the other

hand, these conditions are (mostly) not pathologies, since they are normal within the reference classes in which they (most often) appear.

That different potential individuations of reference classes yield different health judgements should not be considered problematic for Boorse as long as he is right that the relevant reference classes to consider are those individuated by species, sex, and age. However, several authors argue that Boorse is not right about this.

Doust, Walker and Rogers (2017) argue by an example from medicine that it is not clear that medicine always individuates reference classes by species, sex, and age. Their example concerns the US Kidney Foundation's guidelines for classifying kidney disease:

> The new classifications for CKD [chronic kidney disease] use a combination of two measures of kidney function, glomerular filtration rate (GFR), and the presence of albumin in the urine (albuminuria) […]. The thresholds are based on where the risk of cardiovascular disease and end-stage renal failure begins to rise in association with these two measures based on population cohort studies of more than 2 million individuals, including all ages and sexes and a wide range of ethnic groups (Hallan et al., 2012). (Doust, Walker and Rogers 2017, 358)

Also with reference to medicine, Lemoine and Giroux (2016) even question Boorse's usage of the term 'reference class' in the BST, since it is not, they claim, a term used in physiology.

If we understand the BST's definitions of 'health' and 'pathology' as descriptive definitions of the terms as used in theoretical medicine, the above arguments must be regarded as serious criticisms. If we instead understand the definitions as explicative, we may expect some deviation from medical classification and language usage, and these arguments need not be as problematic for the BST.

Yet, if we understand the BST's definitions as more explicative than descriptive, there is another argument against Boorse's account of reference classes. Kingma (2007; 2014) argues that there is no objective ground for choosing a particular definition of reference classes (e.g. involving sexual orientation) over another (e.g. not involving sexual orientation). And without such an objective ground, she concludes, Boorse's choice of definition is value-involving: there are several different definitions to choose between when formulating the theory, and the choice of definition made out of these alternatives is dependent on value judgements rather than objective facts.

Kingma says that what Boorse presumably would answer to her charge is that

> […] reference classes simply *are* the reference classes that are relevant for the distinction between health and disease. Different reference classes would generate different distinctions, but those are not the distinctions between health and disease. Although medicine might have chosen to engage with other distinctions and other concepts, this is only to say that medicine might have concerned itself with things other than health and disease. This does not make the distinction between health and disease evaluative. As he puts the point, '[t]o choose wood over concrete to build your house with is an evaluative choice, but that does not make the concepts of wood and concrete value-laden' (1997:27). (2007, 131, italics are original)

In other words, Boorse may answer to Kingma's charge that the BST's definitions merely describe certain relations in nature.

Despite of this possible answer from Boorse, Kingma argues, the charge about value-involvement is still valid:

> [O]nce reference classes are fixed the BST does not appeal to social judgements to move from the facts about a case to a judgement about its health status. The fixing of reference classes, however, is an evaluative choice which may reflect some deep underlying normative commitments to, for example, ideas about sexual attraction. Therefore I must still conclude that the BST is in all relevant ways evaluative […] (2007, 132)

I think that Kingma is wrong to claim that Boorse's choice of reference classes makes the BST itself value-laden. Even if the choice of reference classes in the BST is motivated by some underlying normative commitment, that does not imply that the theory itself involves any value. This is something that Boorse points out:

> A correct definition of concept H in terms of concepts C_1, C_2, … C_n is value-laden precisely if one of the C_i is value-laden: that is, if a judgement of the form "x is C_i" is a value judgement. It does not matter at all how concept H was "chosen", only what it is […]. (2014, 693)

So, even if the choice of reference classes does reflect some deep underlying normative commitment, that does not imply that the theory itself involves any value. The theory merely describes certain relations in nature.

However, even if Kingma is wrong that a choice of reference classes makes a theory of health and pathology value-laden, one may wonder whether there are values guiding the reference class choice that are nevertheless problematic. That is, one may wonder whether the terms 'health' and 'pathology' should be used to refer to certain relations in nature rather than others (or even whether they should be employed at all). According to Boorse, the choice of reference classes in the BST is not based on value judgements:

> The medical concept of health that I seek to analyze already exists as a target. "Candidate concepts", by contrast, exist only in the minds of philosophers. So the only way to run an argument of this type is to claim that medicine–not the BST–has chosen one of many possible health concepts. The basic problem remains: none of these writers [Kingma and De Vito] give sense to the idea of a value-based "choice of health concept" by medicine. The obvious way to do so fails. That is to assume from the outset that "disease" is an evaluative concept, meaning something like "undesirable condition" or "condition deserving medical treatment". Then medicine will choose what conditions fall under this description–*i.e.*, make value judgements about what conditions are undesirable or need treatment. But, so clarified, the argument has two fatal defects. First, it is circular, since it assumes its conclusion, that "health" is value-laden. Second, it ignores one of the most basic features of medical usage of "disease": that *disease* and *medically treatable condition* do not coincide. As noted, medicine does not call everything it treats a disease, or pathological. Unwanted fertility, unwanted pregnancy, male foreskins, sagging jowls, and small breasts are treated by medicine, yet never counted pathological. (2014, 693, italics are original)

As we can see in this quote, Boorse takes Kingma's argument to rest on the assumption that the BST's individuation of reference classes is based on pre-theoretical judgements of certain conditions as healthy or pathological, and that these pre-theoretical judgements are based on positive and negative evaluations of particular conditions. Boorse is correct to point out that it is questionable whether the BST rests on such evaluative judgements, since the theory does not judge all positively valued conditions healthy or all negatively valued conditions pathological. This is something that also Werkhoven argues:

> First, it is perfectly possible for individuals or cultures to value positively clear cases of pathologies and to value negatively clear cases of health, without this altering the health-status of these conditions. […] Second, there are many physical and mental conditions we tend to disvalue, both personally and culturally,

> that are neither diseases nor health impairments–ugliness, shortness, and stupidity being cases in point. (2020, 152-153)

However, what neither Boorse or Werkhoven acknowledge is that the choice of reference classes may be motivated by other values than positive or negative valuations of certain conditions. Whatever scientists choose to study, there is a value motivating that choice. Although this is perhaps more apparent with regard to certain scientific concepts than others, there is no essential difference between different scientific concepts here. For example, physicists describe different particles and their relations because such knowledge is valued. Now, one may perhaps object that the concepts of protons, electrons, and other particles are different from the concepts of health and pathology in that they reflect natural kinds. However, that should not matter here. Even if physicists are interested in different particles because of them being natural kinds, there is still a value guiding their research – a value attached to natural kinds. So, if the relations described by a theory of health and pathology employing a certain definition of 'reference classes' are interesting, then they are so because of certain values. But these values are not essentially different from values motivating, for instance, physicists' interest for different particles. In response to Kingma, then, this type of value-involvement in the choice of reference classes is not by itself problematic. It is inevitable. Without values guiding researchers in what to study, there would be no scientific knowledge.

Still, though, Kingma's objection is not without merit. Although a theory of health and pathology cannot be accused of being value-laden because of describing certain relations in nature rather than others, it may still be questioned whether physiologists should be interested in the relations that the theory's definitions of 'health' and 'pathology' describe. One may be skeptical towards the values motivating us to consider health and pathology as accounted for by a particular theory. If the choice of reference classes does reflect some morally dubious idea, or is scientifically uninteresting, then one may argue for a change in the definition of 'reference classes'. One may perhaps even argue that no choice of reference classes will make the concepts of health and pathology describe anything interesting and conclude that these concepts should be abandoned.[18]

In the account that I will develop in later chapters, I will try to define 'reference classes' in a way not motivated by remarkable value judgements. Since

[18] For a possible way of defending a choice of reference classes as interesting, see Broadbent (2017). According to Broadbent, we may view health and pathology as secondary properties, arising from our human tendencies to conceive of the world.

physiology is about functioning, I think it makes sense to primarily individuate reference classes by similarities in functional parts. I will suggest a full account of reference classes in Chapter 4 (Section 4.2). But already in Chapter 3 (Section 3.1.4), I will say a bit about what this account assumes.

2.3.2 The Choice of Goals

Much discussion about the biostatistical theory concerns the – so-called – physiological goals. Several authors question whether the relevant physiological goals, and the aspects relevant for health and pathology, are survival and reproduction only, or even survival and reproduction at all. Before considering these specific criticisms, we may remember that Boorse uses the claim that organisms are inherently goal-directed to objectively justify the goals of survival and reproduction (see Section 2.2.2). However, it is questionable whether organisms actually are inherently goal-directed. In the alternative account of health and pathology that I will develop, I will not assume that organisms are inherently goal-directed and hence that there is no objective justification for certain physiological goals. Still, some goals might be more relevant, or interesting, to consider than others. What requires discussion, then, is what goals are relevant to consider.

Schwartz (2007) argues that survival and reproduction are not the only relevant goals. This is because pathological conditions need not reduce survival, or reproduction. For example, aphasia, blindness, or a missing limb need not be associated with reduced chances of survival or reproduction in a society that is suitably supportive (Schwartz 2007, 379). In order to account for all pathological states, Schwartz adds a further aspect to the BST, namely an aspect of negative consequences. So, for a state to be pathological, it need not reduce survival or reproduction. Rather, it is sufficient that it has certain significant negative consequences. Schwartz describes "negative consequences" in the following way:

> the relevant negative consequences should be […] effects that significantly diminish the ability of a part or process in the organism, or of the overall organism, to carry out an activity that is generally standard in the species and has been for a long period of time. (2007, 379)

Examples that Schwartz gives of "standard activities" are speaking, seeing, hearing, and running (Schwartz 2007, 380).

I think Schwartz is right that the BST has problems classifying states that are either given effective treatment (e.g. medical surgery, physiotherapy, psychological therapy), or are effectively compensated for (e.g. by wheelchairs, hearing aids, personal assistance), as pathological. However, the problem of accounting for treatable and compensatable states can be solved (at least to a large extent) without adding an aspect of negative consequences. In the alternative theory that I will develop, I will account for treatable and compensatable conditions as pathological by excluding the possibility of special medical and social interventions when determining survival and reproduction chances (see Chapter 3, Section 3.1.2). However, this solution leaves a version of Schwartz's argument unanswered. One may still question whether some conditions usually judged pathological really do have significant negative effects on survival or reproduction.

Another suggestion of a further relevant goal besides survival and reproduction is mentioned by Hausman (2012). According to Hausman (2012, 537), it is also relevant to consider the instantiation of, or carrying out of, valued traits and activities.

According to Forest and Le Bidan (2016), it is doubtful whether survival and reproduction are relevant goals for physiological functions, health, and pathology at all. Although physiological dysfunctions often coincide with negative effects on survival or reproduction, these consequences are not what make dysfunctions dysfunctions, they argue. To see this, Forest and Le Bidan provide a thought experiment. They invite the reader to consider two worlds, World 1 and World 2. World 1 is our world. World 2 is different from World 1 in one important respect: all beings are exactly like in World 1 except that they are immortal and do not reproduce.

In World 1, Huntington's disease manifest clear cases of pathology. The development of this disease consists of several phases, which include disturbances in the wake-sleep cycle, motor signs, and impairment of mental processes. In a late stage of the disease, the person is unable to walk, talk, eat, and care for herself, and death is usually caused by complications resulting from injuries related to falls, poor nutrition, or choking. In World 2, there are people suffering from a condition that cause the same symptoms as Huntington's disease, except for the life-threatening conditions at the latest stage of the disease. According to Forest and Le Bidan, it is reasonable to consider the Huntington-like condition in World 2 a disease. They conclude:

> In World 2, some conditions would be considered dysfunctional or pathological, even if there is no theoretical background that allows us to say in this case

> that physiological functions are contributions to individual survival and reproduction. […] Consequently, even in such a world as World 2 in which nobody dies or reproduces, we can recognize not only that some states would count as diseases, but approximately *what* kinds of states would count as diseases. This means that maybe Boorse's intuition about a theoretical conception of health and disease based on a conception of normal functions as contributions to survival and reproduction is not the right one. Surely we refer to normal functioning when we consider a state as a disease, but maybe we don't consider normal functioning relatively to survival and reproduction in all cases. (Forest and Le Bidan 2016, 45, italics are original)

Forest and Le Bidan then suggest an alternative account of the goals relevant for physiological functions. I will discuss their suggestion in Chapter 3 (Section 3.1.1).

As indicated by the above arguments, it is questionable whether survival and reproduction are the (only) relevant goals for physiological functions, health, and pathology. In Chapter 3 (Section 3.1.1), I will argue that the goals of survival and reproduction are at least relevant to consider. I will also argue that neither Schwartz's, Hausman's, or Forest and Le Bidan's suggestions of relevant goals or aspects are reasonable to adopt. I will conclude that although it is not precluded that there is some relevant goal besides survival and reproduction, it seems very difficult to formulate such a goal.

A further worry about Boorse's choice of goals, which should be mentioned, is that it rests on value judgements. This is a concern raised by several authors (Agich 1983; Brown 1985; DeVito 2000; Engelhardt 1976; Schaffner 1993). For example, DeVito, although agreeing with Boorse that survival and reproduction are the goals of interest in physiology, takes the choice of these goals to be value-laden: "… the choice of the goals and interests of physiologists is value-laden because physiologists define disease (or pathology) and normality on the basis of their subjective interests" (2000, 542).

This worry about the choice of goals being value-laden relates to the discussion about value-involvement in Chapter 1 (Section 1.1.10), where the justification of a theory, rather than the theory's content itself, is based on values. It is thereby also similar in character to the worry that the choice of reference classes makes the BST value-laden. In Chapter 3 (Section 3.1.1), I will explain how I take this worry into consideration in my account.

2.3.3 Kingma's Dilemma

In the introductory chapter (Section 1.2.6), I held as a desideratum that a theory of health and pathology must account both for (i) the situation-specificity of physiological functions, and (ii) situation-specific pathology. And I noted that Kingma argues that a theory like Boorse's will not be able to do so. Kingma (2010) presents a dilemma for Boorse's theory.[19] The upshot is that the biostatistical theory cannot account both for (i) the situation-specificity of physiological functions and (ii) situation-specific pathology, and that the BST, therefore, is not an adequate account of health and pathology.

Let us consider Kingma's argument in more detail. Kingma argues that Boorse is trapped in a two-horn dilemma. On the one horn, he adopts situation-specific statistics, which makes him able to account for the situation-specificity of physiological functions. This choice, however, makes him unable to account for situation-specific pathology. If a necessary condition for pathology is statistical abnormality, then what is statistically normal in certain situations cannot count as pathological. On the other horn, Boorse instead uses non-situation-specific statistics. This makes him able to account for (at least some) situation-specific pathology (in terms of non-situation-specific statistical abnormality). But this also makes him unable to account for the situation-specificity of physiological functions.

Let us look at the example Kingma (2010, 251-252) uses to illustrate the problem. Kingma invites us to think of a human being on four different occasions. After Hausman (2011), we call her Carol. We are told to focus on Carol's digestive subsystem, and the physiological function of digesting food. The four occasions are: (1) relaxing after a meal, (2) after a period of fasting, (3) in the middle of a lengthy exercise session, and (4) after ingesting a specific poison that immobilizes the digestive system (but has no other effects).

Carol responds typically for her reference class to the different situations. On occasion (1) blood flows to the digestive system, and the digestive system digests at maximal capacity. On occasion (2) little blood flows to the digestive system, and the system is fairly inactive. On occasion (3) the body's blood flow is directed primarily to the skeletal muscles, leaving the digestive system with a very low blood flow and virtually dormant. On occasion (4) little blood flows to the digestive system, and the digestive system is virtually dormant.

Together, occasions (1), (2), and (3) illustrate well the above point that health and pathology are not just a matter of performances of physiological

[19] Earlier and more rudimentary versions of the argument are made by Nordenfelt (1993, 90-91) and van der Steen and Thung (1988, 90).

functions, but of performances of physiological functions related to the situation of the performances. When relaxing after a meal, the healthy response by the digestive system is to digest food. After a period of fasting the healthy response by the digestive system is to be fairly inactive (since there is no, or very little, food to digest). And in the middle of a lengthy exercise session the healthy response by the digestive system is to be virtually dormant. In these three cases, the responses are statistically normal given the situation. On the BST, they are deemed healthy by being normal responses to the situation. In the fourth case, however, the digestive system responds normally to the occasion at hand, yet the digestive system is not considered healthy. This, Kingma points out as a counterexample to the BST.

With regard to the last example, I think we should be more precise about how the poison affects Carol. An important point about this example, as it is discussed both by Kingma and others, is that the correct judgement about Carol's digestive system is that it is pathological. However, depending on how the example is understood in more detail, I think the right judgement about the health status of Carol's digestive system will differ.

I think it is correct to say that the poison makes Carol's digestive system pathological if we understand the example in the following way. The poison affects Carol's digestive system so that the digestive system's disposition for digesting food becomes worse with regard to some physiological goal than what is required for health. For example, if the poison makes Carol's digestive system disposed to not digest in a situation where there is food but no poison in it, then it makes Carol's digestive system pathological.

I think it is incorrect to say that the poison makes Carol's digestive system pathological if we understand the example in the following way. The reason that Carol's digestive system does not digest is that the poison merely affects the situation of the digestive system, and not its disposition to digest. The poison may change the situation of the digestive system, for example, by affecting some trait token in Carol, on which the digestive system is dependent, so that this other trait token becomes disposed to behave in a way that no longer adequately enables the digestive system to digest. Or, the poison may just change the milieu of the digestive system into one where the digestive system is not able to digest.

By some reflection, I think it should be clear that Carol's digestive system should count as pathological only on the first understanding of the example. On the second understanding, there is nothing wrong with Carol's digestive system (given that it functioned completely normal before the ingestion of

poison). We would not try to make Carol better off by intervening on her digestive system. Rather, if the poison obstructs the digestive system from digesting by affecting some other trait token on which the digestive system is dependent, we would intervene on that trait token. Or, if the poison hinders the digestive system from digesting by changing the milieu into one where the digestive system is not able to digest, we would try to remove the poison, not intervene on the digestive system.

Does the BST fall prey to the dilemma formulated by Kingma? Both Hausman (2011) and Boorse (2014, 704-705) argue that it does not. As they correctly point out, when Kingma formulates the dilemma she does not distinguish between how tokens of a trait type are normally affected by situations, and normal dispositions. In the case where Carol has ingested a poison that negatively affects her digestive system's disposition to respond to various situations, the BST actually judges it pathological. This is because the digestive system in this case does not have the disposition to perform all of its trait type's physiological functions on typical occasions with typical efficiency. Although it is normal for digestive systems to be disposed to not digest in the presence of the poison that Carol has ingested, it is not normal for them to be disposed to not digest when there is food but no poison in the digestive system. Whether or not Carol's digestive system is in a situation where it is typical to acquire the inability that Carol's digestive system acquires in the poisonous situation is irrelevant. As long as it is generally not typical for digestive systems in Carol's reference class to not digest food when there is food but no poison in the digestive system, Carol's digestive system does not have the disposition to digest with typical efficiency on typical occasions for digestion.

So, there is no problem for the BST to account for Carol's digestive system as pathological when the poison changes the digestive system's disposition for digesting to be less beneficial for some physiological goal than what is required for health. However, the BST is not able to say that Carol's digestive system is pathological if the poison only hinders the digestive system from digesting as long as it is present in the body, since it is typical for digestive systems to not digest in that situation. In that case, Carol's digestive system is not differently disposed than most other digestive systems. About this sort of cases, Boorse writes:

> The only kind of disease to which her [Kingma's] argument in the quotation is relevant is diseases maintained by an environmental factor, with no lasting change in the organism. This is a relatively tiny class, of conditions such as heat exhaustion, mountain sickness, or nitrous-oxide intoxication, which are almost instantly cured by change of environment. (2014, 705)

I think that what the BST implies with regard to this "relatively tiny class of conditions" is correct. If it is typical in the reference class to be disposed to not digest when there is poison in the digestive system, then there is nothing wrong with Carol's digestive system if it merely does not digest in the presence of the poison. There is nothing to improve about the digestive system. Carol's digestive system is pathological only if it acquires an untypical inability, for example not being able to digest when there is food but no poison in it. Importantly though, the conclusion is not that Carol's digestive system can only be pathological when the poison has left her body. If her digestive system is disposed to not digest in a situation where there is food and no poison in it, then the digestive system is pathological, regardless of the current situation.

Kingma (2016) thinks that Hausman's answer to her objection (which I just argued solves the issue) does not solve the problem. She writes that Hausman's response "which consists in comparing normal dispositional function against the whole population or reference class, rather than against organisms in similar circumstances" (Kingma 2016, 392) "fails to generalize and thus does not adequately answer my [Kingma's] ([2010]) objection" (Kingma 2016, 392). Kingma argues that Hausman's response fails to generalize by a number of examples. I will argue that Kingma's examples do not rebut Hausman's response. This is because, in these examples, Kingma does not carefully consider what is part of the situation, or the circumstances, of the trait token under evaluation. In order to see this, let us consider three of Kingma's examples. The first is about pregnancy:

> Take, for example, a healthy pregnant woman–that is, a woman who performs statistically typical functions for being at the stage of pregnancy that she is in. According to Hausman, the normal dispositional function of such a person should be compared to that of the population or reference class as a whole rather than, on Kingma's ([2010]) interpretation, to that of other pregnant women. That commits Hausman to recognizing a slew of pregnancy-related pathologies: amongst many others, pregnant women–in comparison to the whole population–have a reduced ability to run, lift, or bend over; are more disposed to sustain joint injuries because hormonal changes make soft tissue more flexible; have some suppression of the immune system; are less able to store moderate quantities of urine in their bladder; and lack various other normal physical abilities, such as the ability to directly come up to a sitting position from a position of lying on one's back, or the ability to keep sufficient tension in one's pelvic floor whilst sneezing or coughing. Against Hausman, however, none of these problems are medically or conventionally considered pathologies; they are considered a normal aspect of being pregnant. (2016, 395)

Here, Kingma's conclusion that Hausman's response classifies the conditions she mentions as pathological is mistaken. Kingma is right that Hausman claims that when we evaluate the health status of some particular trait token of a pregnant woman, we should compare this trait token's dispositional function against the whole reference class, rather than merely against other pregnant members of the reference class. However, Kingma is wrong that a healthy pregnant woman has different dispositional functions compared to other non-pregnant women in the same reference class simply by being pregnant. If we, for example, consider the immune system of a healthy woman, it does not become differently disposed when a woman gets pregnant. The immune system will not perform its type's physiological function with as high intensity during the woman's pregnancy as otherwise; however, this is not because it gets differently disposed during the pregnancy, but because the situation imposed by the pregnancy makes the immune system perform less intensely. This lower intensity in situations of pregnancy is in accordance with how healthy immune systems are disposed to carry out their physiological function. To be clear, what Kingma misses to take into account in this example is that the situation, or the circumstances, inter alia include the facts that there is a fetus in the uterus, and that certain hormones are at certain levels. With regard to such situations, immune systems of non-pregnant women are not different from immune systems of pregnant women: non-pregnant women's immune systems are also disposed to respond with lower intensity in situations where there is a fetus in the uterus and certain hormones are at certain levels. However, in a case where the pregnancy actually changes the immune system's disposition to the worse, for example so that it becomes disposed to respond with a lower intensity also when hormone levels and other circumstances are back to that of a non-pregnant situation, the immune system will count as pathological.

Another of Kingma's examples concerns breastfeeding. Kingma writes:

> Hormonal changes involved in lactation suppress most women's menstrual cycle and ability to conceive. It also diminishes a woman's ability to lubricate the vaginal wall in response to arousal (Alder [1989]). Both effects are or can be considered pathological in non-breastfeeding women, but are considered normal and healthy for breast-feeding. (2016, 395)

Just like in the previous example, Kingma here misses to take into account that breastfeeding is part of the situation. Hence, there is no difference be-

tween breastfeeding women and non-breastfeeding women in the dispositional functioning of their menstrual cycle, or lubrication of the vaginal wall. Both breastfeeding and non-breastfeeding women are functionally disposed to not menstruate and to have a reduced lubrication of the vaginal wall in response to arousal in circumstances where they are breastfeeding.

A further example, which does not involve reproductive functions, concerns sleep:

> […] take an organism that is asleep (and preferably an organism that is not sleeping most of the time, so a human rather than a lion). Such an organism lacks many of the dispositional abilities that a waking organism has. It is suffering from paralysis, loss of consciousness, and it is not disposed to respond to very many sounds and sights that it would ordinarily respond to, and that it is normal to respond to. On Hausman's interpretation of the BST, that would make sleeping a pathological condition. (Kingma 2016, 396)

Also here Kingma is mistaken in her conclusion because she does not pay attention to the different situations of sleeping and awake organisms of a reference class. Although sleeping organisms and awake organisms behave rather differently, they are all disposed to suffer from paralysis, loss of consciousness etc. while in a situation of sleep (e.g. when melatonin levels are high).

Now, one may perhaps object against my response that internal factors such as having a fetus in one's uterus, or having certain hormone levels, cannot be part of the situation – and rather hold that the situation only includes external factors. However, I think that the only reasonable view here is to count internal factors as part of the situation. Remember that one important part of Kingma's dilemma was that an account of health and pathology must be able to account for the situation-specificity of physiological functions. For many physiological functions, what is an adequate performance varies with internal facts. Consider, for example, the presence of influenza viruses within one's respiratory system which makes it adequate for the immune system to become active. Or consider the presence of food within one's mouth which makes it adequate for the salivary glands to produce saliva, or the presence of food within one's stomach which makes it adequate for certain cells in the stomach to release gastrin, whose presence in turn makes it adequate for certain other cells in the stomach to release hydrochloride acid. Also if we look at Kingma's own examples, it should be clear that the situation must include not only external, but internal factors. In order to account for the health of a suppressed immune

system in cases of pregnancy, the immune system's reactivity must be relativized to internal factors, for instance that there is a fetus in the uterus, or that certain hormones are of a certain level. Similarly, in order to account for the health of an unconscious brain in cases of sleep, the brain's consciousness must be relativized to internal factors, for instance the level of melatonin.

The reasonable view to take, given the above arguments, is that the performance by a trait token is relative to everything, both outside and inside the organism. When I start to develop my alternative theory of health and pathology in the next chapter, I will use this idea explicitly (see Section 3.2.3). Although Hausman's response to Kingma is basically correct, he does not discuss situations in detail. By making it clear what situations include, it should be apparent that Kingma's dilemma is not a genuine dilemma: the BST may account both for situation-specific diseases and the situation-specificity of physiological functions.

There is, however, an argument to make against the BST, which is related to Kingma's dilemma. The BST does not, for another reason than what Kingma brings up in her dilemma, fully account for the situation-specificity of physiological functions. In the BST's definition of 'health' and 'pathology', it is only taken into consideration how the trait token is disposed to perform its type's physiological functions *on typical occasions*. However, what is relevant for health is not only the ability to perform certain physiological functions in situations in which it is adequate to perform them. In order to be healthy it is also often important to not perform physiological functions in situations in which it is inadequate to perform them. To see this, consider, for example, autoimmune diseases where the immune system performs its physiological function also in situations when there is no threat to incapacitate. Or, to relate to the Carol case, it should not be considered healthy to produce a lot of gastric juice when there is no food in the digestive system. In the alternative account that I will present, not only situations in which performances of a physiological function are appropriate or typical, but also situations in which performances of a physiological function are inappropriate or atypical, will be taken into account when determining the health status of a trait token.

2.3.4 The Problem of Common Diseases

In the introductory chapter (Section 1.2.6) I brought up as a desideratum that a theory of health and pathology should be able to classify states of common diseases as pathological. This is, as pointed out by Schwartz (2007) and

Hausman (2012), a type of condition that the biostatistical theory is not able to account for as pathological. Let us see why.

As we saw, the BST's exact line distinguishing health from pathology is conventionally drawn. But the line should at least be drawn statistically and be located below, or far below, the average efficiency level. These requirements allow for several definitions of the distinguishing line. Let us be more precise. In accordance with the BST, trait token performances of physiological functions are ascribed efficiency levels. Let us denote the function making these ascriptions e:

e : performances of physiological functions $\rightarrow$ efficiency levels

such that $e(p)$ is the efficiency level of the performance p.

Although Boorse does not say exactly what efficiency levels are, he at least assumes that they are linearly ordered and are such that we can define statistical measures such as averages and standard deviations. This means that efficiency levels could be, for example, real numbers, non-negative real numbers, or real numbers in the interval 0-1. Let us, for the purpose of some generality, assume that efficiency levels are real numbers.

As explained in Section 2.2.4, for the distinguishing line to count as statistically determined, there must be some not too "exotic" function

d : distributions of efficiency levels $\rightarrow$ efficiency levels

such that a trait token is pathological, relative to a reference class R, if and only if $e(p) < d(E)$ for at least one of its performances p of a physiological function, where E is the distribution of efficiency levels of the trait token performances of that physiological function in R. Some definitions of d, which can reasonably be claimed to draw the line statistically, are the following (where $avg(E)$ and $std(E)$ denote the average and the standard deviations of the distribution E, respectively):

$d(E) := avg(E) - c$ (c a constant real number, $c > 0$)

$d(E) := c \times avg(E)$ (c a constant real number, $0 < c < 1$, $E \geq 0$)

$d(E) := avg(E) - c \times std(E)$ (c a constant real number, $c > 0$)

Importantly, the BST cannot allow for different definitions of d in different cases, and particularly not different values of c in different cases, since that would mean that the distinguishing line is not statistically determined but set by something else. But none of the three possible interpretations of d above will provide reasonable results both for tokens of trait types affected by common diseases and tokens of trait types where pathology is rare. And it seems very unlikely that one could come up with some other not too exotic definition of d that does so.

In Section 2.2.4, I mentioned that one perhaps may interpret the BST's drawing of the line distinguishing health from pathology in a weaker way. According to this interpretation, the line would not necessarily be statistically determined. However, the line would still need to end up below, or far below, the average efficiency level. While this interpretation leaves it a relatively open question how to define the distinguishing line, it allows for larger variations in frequency of pathology between different trait types. But still, since the line must end up below, or far below, the average efficiency level, it is questionable whether it allows for 30 percent of the trait tokens of a type being pathological. Moreover, there may be pathologies affecting 50 percent or more of the trait tokens, which would be even harder to account for.

The alternative account of health and pathology that I will develop in Chapters 3-5 will account for common diseases, partly by defining the standard for health differently from Boorse, and partly by drawing the line between health and pathology in another way than Boorse does.

2.3.5 Further Objections

I have discussed four types of objections to the biostatistical theory, objections concerning the BST's account of reference classes and its choice of physiological goals, Kingma's dilemma, and the problem of common diseases. I will come back to these objections in later chapters.

It may be mentioned, though, that the BST has received a vast number of further objections. Boorse has discussed and responded to most of the criticisms posed up to 2014, mainly in the *A rebuttal on health* (1997), *A rebuttal on functions* (2002), and *A second rebuttal on health* (2014), but also in *Disability and medical theory* (2010).

One type of objection that the BST has received, which may be mentioned here, concerns value-involvement. Remember that I held as a desideratum that the theory that I am interested in developing should be value-free. This means that objections about value-involvement are as relevant to the theory that I will

develop as to the BST. Besides the already discussed objections about value-involvement concerning the choice of physiological goals and concerning the choice of reference classes, there are, first, further objections claiming that the BST is not value-free as Boorse claims. Reasons given for this claim are that: science is value-laden (Agich 1983), biology is a value-laden science (Agich 1983), anatomical and physiological descriptions are value-laden (Stempsey 2000), and specific terms of the BST are value-laden (DeVito 2000; Fulford 1989). Second, there are objections aiming to show that the concepts of health and disease, in contrast to what Boorse thinks, must be value-laden. Several authors argue that practical aspects must matter, since medicine essentially or primarily is a practical activity or an art (Agich 1983; Brown 1985). Several authors also argue by examples that values must play a role in the determination of the health status of a state (DeVito 2000; Hare 1986; Merskey 1986).

With regard to all of these further objections, however, I think that Boorse (1997; 2014) has provided satisfying answers. I will hence not discuss them further in this thesis.

3 Introducing the Disposition Profile Efficiency Theory

I will now start to develop my theory of health and pathology. I call it 'the disposition profile efficiency theory', or abbreviated 'the DPE-theory'. The theory is similar to Boorse's in that it defines 'health' and 'disease' with reference to reference classes, physiological functions, efficiency, and a line-drawing formula. Yet the theory is different in more detailed aspects that are important for fulfilling the desiderata from Chapter 1, with which the biostatistical theory struggles. The disposition profile efficiency theory is more fully articulated than the BST, and hence clearer. This can, to start with, be seen in the theory's core part, namely models of dispositions for performing physiological functions. It can also be seen in the DPE-theory's definitions of 'reference class', 'physiological function', 'efficiency', and the line drawing formula, which all differ from the BST. The models of dispositions for performing physiological functions, together with the definition of 'efficiency', makes it possible to account for the situation-specificity of physiological functions. These parts of the theory are also crucial for accounting for health and pathology as reference class-relative properties, without disregarding the importance of integration for health. Further, they help making it clear that Kingma's dilemma is no genuine dilemma. The definition of 'efficiency' and the line-drawing formula makes it possible to account for common diseases. And, it should be clear from the DPE-theory's accounts of reference classes and physiological functions that the DPE-theory is not value-involving.

The theory will be presented in detail through Chapters 3, 4, and 5. To start with, I will in this chapter first present basic assumptions of the DPE-theory (Section 3.1). Then, I will account for the most basic part of the theory – disposition profiles (Section 3.2).

3.1 Basic Assumptions

Beneath, I will present basic some assumptions of the DPE-theory. These concern the goals relevant for physiological functions (Section 3.1.1) and how to measure these goals (Section 3.1.2), how to handle several goals when measuring efficiencies and health statuses (Section 3.1.3), and, lastly, how determinations of reference classes, trait types and physiological functions depend on each other (Section 3.1.4).

3.1.1 Relevant Goals

Like Boorse, I assume a goal-analysis of physiological function. In response to the criticisms brought up in the last chapter (Section 2.3.2), I will discuss the goals that are relevant for physiological functions, and the justification for these goals. As will become clear, my views on physiological functions differ somewhat from Boorse's.

Let us start with the question of justification. We saw that Boorse thinks that organisms are goal-directed. This goal-directedness, Boorse argues, makes the goals of survival and reproduction objectively justified. They are objectively justified by the fact that organisms, as a fact of nature, strive towards them. However, the view about organisms being goal-directed is questionable. Hence I will not assume it. Without such an objective justification for certain goals, physiological functions, health, and pathology cannot reasonably be claimed to be natural kinds. But, as explained in Chapter 1 (Section 1.1.9), this does not exclude that physiological functions, health, and pathology are natural properties. Rather, according to the DPE-theory, they are. Although these notions do not necessarily carve nature at its joints, they describe potentially interesting relations between more basic natural properties (possibly natural kinds).

A certain choice of goals does not make the DPE-theory value-laden. By saying that a state is healthy or pathological, one only makes a descriptive claim about the relation between the state (of a trait token in an individual of some reference class) and a certain goal; there is no claim that this goal has another status than other potential goals. In principle, we could choose to consider any naturalistically describable goal. However, some goals are reasonably more relevant, or interesting, to consider than others. The relevant goals to consider, I take it, are goals that are of interest in physiology and pathology, and that do not make the theory deviate too much from very basic ideas of physiological functions, health and pathology in these disciplines.

Which are these goals? We saw in the last chapter that several authors question Boorse's claim that survival and reproduction are the (only) relevant goals for physiological functions. Forest and Le Bidan (2016) indicate that survival and reproduction are not even relevant goals for the determination of physiological functions. According to Forest and Le Bidan, dysfunctions often coincide with negative effects on survival and reproduction, but these effects are not what makes dysfunctions dysfunctions. Recall their thought experiment meant to show this, about Huntington's disease in World 1 (our world) and a Huntington-like condition in World 2 in which people are immortal and unable to reproduce.

Survival and reproduction, I think, should at least be considered relevant goals. States that in principle have no other effects than significantly reducing survival or reproduction are clearly considered pathological. Think for example of a state in the reproductive organs that does not have any other consequences than hindering fertilization. Or think of a brain tumor that does not cause any other significant symptoms other than death, except perhaps for pain in the last days of life. Although there is pain for a rather short period of time, that is not what makes the state a serious pathology. What makes the state a serious pathology is rather the fact that the tumor rapidly causes death.

One may perhaps question whether reproduction is a relevant goal, since it is not normal for all reference classes that their members reproduce. Some species do not reproduce at all (e.g. mules), and in some species reproduction does not occur among older individuals (e.g. human females). However, this should not be taken to imply that reproduction is not a relevant goal for physiological functions, health, and pathology. If mules and elderly human females became fertile, their reproduction would be relevant to consider.

Are there other relevant goals besides survival and reproduction? Schwartz (2007) considers survival and reproduction as relevant goals; however, he thinks that these goals are not sufficient to account for all pathology. As explained in Chapter 2 (Section 2.3.2), Schwartz thinks that the goals of survival and reproduction are not enough to account for conditions like aphasia, blindness, and a missing limb as pathological, since these conditions do not affect survival or reproduction in a society that is suitably supportive. As I responded in Chapter 2, many treatable and compensatable states can be accounted for as pathological if we exclude the possibility of special medical and social interventions when determining survival or reproduction chances. (I will come back to this in the next section.) However, even if we may account for the pathology of states that have significant negative effects on survival or repro-

duction without special medical and social interventions, one may still question whether some conditions usually judged pathological really do have significant negative effects on survival or reproduction even if special medical and social interventions are excluded. Let us, therefore, consider potential additional goals suggested by Schwartz, Hausman, and Forest and Le Bidan.

In accordance with Schwartz's (2007) suggestion, a further goal would be the carrying out of activities that are standard in the species and have been for a long period of time. Examples of such activities, mentioned by Schwartz, are speaking, seeing, hearing, and running (2007, 380).

Schwartz's suggestion is problematic, I think, because it ascribes physiological functions to activities that are clearly not recognized as physiological functions in physiology. It ascribes physiological functions, not only to activities such as speaking, seeing, hearing and running, but also to activities such as experiencing period pain, expressing premenstrual syndrome symptoms, or showing symptoms of standard diseases (e.g. colds, chicken-pox). These are activities that are standard in the species and have been for a long period of time. On Schwartz's suggestion, it may even be pathological to be unable to experience period pain, to express premenstrual syndrome symptoms, or to show itchy rashes during an infection by chicken pox.[20]

A somewhat similar suggestion of a further goal by Hausman (2012, 537) is the instantiation of, or carrying out of, valued traits and activities.[21] This suggestion does not have the same problem as Schwartz's. Although activities such as being in pain during one's periods, or showing itchy rashes during an infection by chicken-pox are, and have been, standard for humans, they are not valued.

It is not clear whether the suggestion is that the traits and activities are valued by the organism at stake, or whether they are generally valued among individuals in the organism's reference class. Irrespective of which of these ways we understand it, though, I will argue that the suggestion is problematic. This is because physiological functions are not, I think, determined by the values of particular individuals, or social values.

---

[20] See Hausman (2012, 529) for further arguments against Schwartz's suggestion.

[21] It is, however, questionable whether Hausman actually considers this as a further relevant goal, since he also writes: "What distinguishes defect from health is not whether the condition is good for a person or other organism […] but whether the condition contributes to the functioning of some directively organized system within the organism, where that system directly or indirectly typically promotes fitness" (2012, 520). Since I am here interested in the question whether there are any further relevant goals besides survival and reproduction, it is interesting to consider this suggested goal in any case.

If physiological functions were determined by the values held by particular individuals, then physiological functions would not always be associated with types of traits in a reference class. Rather, we would have to investigate what the physiological functions in an individual are by considering the values held by it. This does not only sound strange and distant from physiological theory, it also violates the desideratum about reference class-relativity: if there are no classes of organism sharing physiological functions, then health and pathology cannot be reference class-relative properties.

If physiological functions were determined by values generally held among individuals in a reference class, then the physiological functions would not differ between different individuals of a reference class. However, this interpretation is problematic since it makes ascriptions of physiological functions directly dependent on cultural norms and ideals. Even if circumcised female genitals or very small feet in females were widely held norms among a reference class of women, this should reasonably not mean that circumcised female genitals, or very small feet, serve physiological functions. Further, it should reasonably not mean that non-circumcised female genital organs, or non-bound feet, are pathological. What could instead be said about such cases is that circumcised female genitals or very small feet in females serve *social* functions, and that non-circumcised female genital organs, or non-bound feet fail to serve these.

What about Forest and Le Bidan's suggestion, then? According to Forest and Le Bidan, the relevant goals for physiological functions are capacities that "share three important features" (2016, 46). The first feature is that the capacity is distinctive for the reference class. It should, for example, not be shared with non-living beings (e.g. the capacity of one's body to make a shadow on the wall). The second feature is that the capacity "result[s] from one's psychological and physiological organization, not from sociological or political conditions" (Forest and Le Bidan 2016, 46). Examples of such capacities, according to Forest and Le Bidan, are capacities to read, count and memorize; however not the capacity to vote in a presidential election. The third feature is that the capacity is typically needed for usual interactions between a member of a reference class and her environment. "For instance", Forest and Le Bidan write, "the function of the hip is to join the thigh to the pelvis, and ultimately, to make it possible for us to walk" (2016, 46-47).

For a similar reason that Schwartz's added aspect of standard activities should not count as a relevant goal, I think that Forest and Le Bidan's suggestion is not a reasonable account of relevant goals. Forest and Le Bidan, at least theoretically, limit the class of physiological functions in comparison to

Schwartz by including, not all standard activities, but only those that can be described as interactions between an organism and its environment. However, I will argue, this class still includes activities that are typically not considered physiological functions. Just as well as walking can be described as an interaction between an organism and its environment, so can common symptoms of periods such as cold sweats and diarrhea. Just as walking is an interaction between the body and the ground and the air, cold sweats and diarrhea are ways in which the body interacts with the near environment (e.g. the air, or a water closet). By similar reasoning, usual interactions between human females in fertile ages and their environment include acting in accordance with symptoms of premenstrual syndrome. And among young humans, itching and showing red rashes are usual interactions with the environment when having chicken pox. Hence, because of Forest and Le Bidan's rather widely defined third feature, capacities to cold sweat and have diarrhea during one's periods, to act in accordance with symptoms of premenstrual syndrome, and to itch and show red rashes when having chicken pox come out as relevant goals for physiological functions. Forest and Le Bidan's account then also implies that not having the capacity to cold sweat and have diarrhea during one's period, or not having the capacity to show red rashes when infected by chicken-pox, may be pathological.

To sum up the attempts of formulating some additional goal besides survival and reproduction, Schwartz's and Forest and Le Bidan's suggestions fail because not every standard activity, or distinct, typical function arising from the organism's biological construction is considered a physiological function. Hausman limits the range of physiological functions by a requirement of being valued. This, however, inadequately ascribes physiological functions by cultural norms and ideals.

I do not want to exclude the possibility of some further relevant goal besides survival and reproduction. However, the best we can do at the moment, I think, is to go with only the goals of survival and reproduction. In the end, what makes capacities such as walking, hearing, and writing physiological functions is perhaps merely the fact that they ultimately contribute to survival or reproduction, although some only slightly.

3.1.2 Measuring the Goals

I just concluded that physiological functions, according to the disposition profile efficiency theory, are contributions to the goals of survival or reproduction, just like on the biostatistical theory. How do we measure how successful

an organism is with respect to survival and reproduction? This is something that Boorse does not discuss. A reasonable way to measure survival chances, I think, is by further life expectancy, i.e. the time that the organism is expected to continue living past the time of measurement. So, an organism with high survival chances is an organism whose further life expectancy is high, relative to its reference class.

Why should we measure survival chances by further life expectancy, rather than simply by life expectancy? Although this is connected to aspects of the DPE-theory that will be discussed first in later chapters, let me try to explain why. Both what is a typical further life expectancy and what is a typical life expectancy will vary during an individual's life. The life expectancy generally increases the older the organism gets, since a certain age is already guaranteed, whereas the further life expectancy generally decreases the older the organism gets, since less of life is left. As I will come back to in Chapter 4 (Section 4.2), this means that reference classes need to be individuated by age. On the DPE-theory, when we evaluate the health status of a trait token in an organism, what we do, in very rough terms, is to compare the organism's survival chances to the survival chances of a standard organism for the reference class. We consider the ratio between the survival chances of the organism with the trait token under evaluation and the survival chances of the standard organism (this will be carefully accounted for in Chapters 4-5). And here we find the reason why we should measure survival chances by further life expectancy rather than by life expectancy. If we measure survival chances by life expectancy, then it will in principle be impossible for individuals belonging to reference classes of very old organisms to have pathology: in percentual terms, the life expectancy of the standard will be very close to the life expectancy of the individual under consideration, even if the trait token we evaluate functions so badly that the organism will die any minute. For example, for a reference class of elderly humans, the comparison may perhaps be between 100 years (for the individual under consideration) and 101 years (for the standard). However, if we instead measure survival chances by further life expectancy, the percentual difference between an old organism with a trait token functioning so badly that the organism will die at any minute and the standard organism will be larger. For a reference class of elderly humans the comparison may perhaps be between 1 day and 1 year.

What about reproduction chances? The best way to measure this calls for more discussion. For example, should the measurement be defined in terms of number of offspring, or number of viable offspring, or spreading of genes? Perhaps physiology and pathology would benefit from several different health

and pathology concepts here, for example one concerning the goal of individual offspring and another concerning the goal of spreading one's genes. For the purposes of this thesis, I will measure chances of succeeding in reproduction by further expected number of offspring. But other different reproduction-health concepts could be construed along the same lines.

Here we may also ask why we should measure reproduction chances by further expected number of offspring rather than by expected number of offspring. Since I will not suggest that we individuate reference classes by present number of offspring, the reason for including "further" here is different from the reason we should measure survival chances by further life expectancy. The reason why we should consider further expected number of offspring, rather than expected number of offspring, is that health, I think, should concern the organism's ability to reproduce in the present and the future, but not its past ability to reproduce.

Connecting to the desideratum that the theory should account for health and pathology as reference class-relative properties, there is an important qualification to make about the measurements in terms of further life expectancy and further expected number of offspring. The determination of an organism's further life expectancy, or its further expected number of offspring, must be relative to an idea about the future. However, in order to account for health and pathology as reference class-relative properties, the determination of an individual's further life expectancy, or further expected number of offspring, should not be relative to the expected future situations of the particular organism. This is because a particular organism may be in a very lucky or unlucky situation with regard to further life expectancy, or further expected number of offspring. And this should not affect its trait tokens' health statuses. For example, an emperor penguin who has lost its flock has worse chances of surviving when it gets very cold. However, this should not mean that this penguin's skin, for example, is pathological. Whether it is or not should rather be relative to a distribution of probabilities representative for the penguin's reference class. In contrast to the penguin who has lost its flock, most emperor penguins are able to huddle together in very large groups when it gets cold. Hence, the further life expectancy of any emperor penguin should be determined in relation to an assumption that it is likely that there is a large group of other penguins to huddle together with when it gets cold. Because of this, I will use the terms 'further life expectancy' and 'further expected number of offspring' in a way somewhat different from the most intuitive understanding of them. As I will use them, the further life expectancy, and the further expected number of offspring, of an individual will be relative to a distribution

of probabilities, assigned to various possible future situations, representative for the individual's reference class.

A further important qualification concerns medical and social interventions. When discussing the BST in the previous chapter, I said that the BST cannot account for the pathology of states that are effectively treated or effectively compensated for. In order to handle treatable and compensatable pathology, we must, when determining further life expectancy and further expected number of offspring, disregard potential special medically or socially directed interventions. By "special medically or socially directed interventions", I include all interventions that are not standardly offered, but only specially offered to certain people (usually those in need). It will, for example, include medical surgery, physiotherapy, psychological therapy, usage of wheelchairs or hearing aids, and personal assistance.

Disregarding special medically or socially directed interventions should not be considered ad hoc. When we determine the health status of a trait token, we do not want to know whether the state reduces life expectancy if treated or compensated for. We want to know whether it reduces life expectancy if not treated or compensated for.

It should also be pointed out that I take the further life expectancy of an organism to be based on complete physiological knowledge of that organism. This is because, on the DPE-theory, health and pathology are not relative to our state of knowledge. Rather, the health status of a trait token in an organism is the same irrespective of what we know about it. For example, if we learn something new about how emperor penguin skin reacts to cold, this will not affect the further life expectancy of the token emperor penguin in the above example. Mutatis mutandis for further expected number of offspring.

3.1.3 Treating the Goals Separately

As stated above, I assume the relevant goals for physiological functions to be survival and reproduction. But in contrast to the biostatistical theory, which has a single concept of efficiency covering both survival and reproduction, I think that the two goals should be given separate efficiency measures. This is, as I argued in the last chapter (Section 2.2.3), because some performances enhance the possessor's chances of survival in comparison to the reference class more than it enhances the possessor's chances of reproduction in comparison to the reference class (or vice versa).

My proposal is to distinguish between survival and reproduction as goals giving rise to two distinct efficiency measures. The suggestion is to talk about

survival-efficiency on the one hand, and reproduction-efficiency on the other. We should note here that if there are several different measures of reproduction chances that are all relevant for physiology and pathology, then there are also several different reproduction-efficiency concepts, for example both a concept of further offspring-efficiency and a concept of gene spreading-efficiency.

With regard to the concepts of health and pathology, we have two options. Either we define 'pathology' as reduction in survival-efficiency or reproduction-efficiency, and 'health' as absence of reduction in survival-efficiency and absence of reduction in reproduction-efficiency. Recall the example in Chapter 2 (Section 2.2.3) about the human female, who because of her pituitary's abnormal hormone secretion ovulates very infrequently. The abnormal hormone secretion makes her worse off with regard to reproduction, relative to her reference class, but not worse off with regard to survival, relative to her reference class. With this first option, we get the result that the pituitary is pathological, since its reproduction-efficiency is reduced. Alternatively, we also distinguish between two health and pathology concepts, one for each efficiency measure: reproductive health and pathology on the one hand, and health and pathology pertaining to survival on the other. If we adopt the latter approach, in the above example we get the result that the pituitary is reproductively pathological but healthy with regard to survival. Which of these two alternatives should we choose? In medical theory, the terms usually used are 'health' and 'pathology', not 'survival-health' and 'survival-pathology'. However, if one of the alternatives is more theoretically fruitful than the other, that is a reason to favor that alternative. Since I think it makes physiological theory clearer and more transparent if one distinguishes between different health and pathology concepts based on the different efficiency concepts, I will opt for the second alternative and distinguish between health and pathology pertaining to survival and health and pathology pertaining to reproduction. Note that if there are several different reproduction-efficiency concepts, we have more than two health and pathology concepts. Instead of health and pathology pertaining to reproduction we, for example, have health and pathology pertaining to number of offspring, and health and pathology pertaining to gene spreading.

3.1.4 Reference Classes, Trait Types, and Physiological Functions

It is questionable whether it is possible to satisfactorily account for reference classes, trait types, and physiological functions in a reductionist way, i.e. so

that some of the three entities is determined by the others. Therefore, I will assume that reference classes, physiological functions, and trait types are determined simultaneously. How this is supposed to work out will become clear in Chapter 4 (Section 4.2), where I present an account of reference classes. The reason I present this account only later in the thesis is that it is dependent on other parts of the DPE-theory. Hence, it makes sense to present it when more of the DPE-theory's machinery have been developed. However, let us already now note some claims that the DPE-theory makes about reference classes. The motivation for these claims stems, first, from basic interests in physiology and, second, from theoretical requirements imposed by the role reference classes are supposed to play in the DPE-theory.

First, every reference class contains at least some minimal number of members. There cannot be any reference class consisting of merely one or a few organisms. Second, for each reference class, there is a set of organism part types that are shared by a significant share of the reference class members. These are the trait types in the reference class. And, third, for each trait type there is one or several function types which are also shared by a significant portion of the reference class members. These are the physiological functions in the reference class. Fourth, it holds for each reference class that all members of it are of the same age. Fifth, reference classes both include individuals alive and individuals of a number of generations back in time. Sixth, all members of a specific reference class must have a close enough common ancestor. This means that two individuals separated in time by a large number of generations cannot belong to the same reference class. It also means that if two groups of individuals have not had a common ancestor within a certain number of generations, then they do not belong to the same reference class.

In the last chapter, I discussed objections against Boorse's account of reference classes. We saw that Kingma argues that Boorse's choice of reference classes makes the biostatistical theory value-laden, since there is no objectively justifiable choice of reference classes. I responded that the choice of reference classes does not make the BST value-laden, but that one may question what values guide the interest for studying certain relations in nature referred to by 'health' and 'pathology'. In Chapter 4 (Section 4.2), I will discuss the DPE-theory's account of reference classes in relation to the worry about value-involvement.

3.2 Dispositions to Perform Physiological Functions

On the disposition profile efficiency theory, the health status of a trait token is determined by the efficiency of the trait token's dispositions for performing its trait type's physiological functions. In this section, I will introduce a model of such dispositions.

3.2.1 Feature Types and Feature Values

In line with Kingma (2010), I will assume that physiological functions are individuated by quality, and not by quantity. This assumption implies that in order to evaluate health statuses, we need to be able to account for quantitative differences between performances of physiological functions. This is for two reasons. First, different situations require different responses of a healthy trait token. Second, we want to be able to compare how well different trait tokens – given a certain situation – perform a certain physiological function.

A way to account for quantitative differences is to consider types of features relevant for describing performances of physiological functions. Since health status evaluations, on the disposition profile efficiency theory, concern further life expectancy and further expected number of offspring, the feature types that count as relevant for describing performances of physiological functions are feature types that affect further life expectancy or further expected number of offspring. For the human heart's blood pumping, for example, feature types such as beating frequency and contraction force (I assume) affect further life expectancy. Hence the quantitative description of a heart token's blood pumping will include quantitative values for beating frequency and contraction force. What are relevant feature types to consider will of course differ between different physiological functions. Whereas performances of some physiological functions include contraction, others, for example, include secretion or digestion.

For every physiological function, I assume that there is a set of feature types that matter for further life expectancy or for further expected number of offspring. Call these feature types 'relevant feature types'. I will use $RFT(F)$ to denote the set of relevant feature types for the physiological function F.

Performances of physiological functions can be quantitively described by what I will call 'feature values', i.e. quantitative values for the relevant feature types. For example, if we assume that beating frequency and contraction force are the relevant feature types for the heart's blood pumping, we can describe

a particular performance of a token heart in a certain situation by its contraction force and its beating frequency in that situation. Also, we may compare two heart tokens' blood pumping by comparing their respective beating frequencies and contraction forces for a certain situation.

By considering relevant feature types, we are able to compare all tokens of a trait type, even tokens that we intuitively would say do not perform the physiological function to any degree, for instance hearts that do not beat at all. In the framework of the DPE-theory, we may say that such tokens perform their type's physiological function with feature values indicating no activity. For example, a non-beating heart has the beating frequency 0 and the contraction force 0. This enables us to compare the efficiency of every token of a trait type in a reference class.

A question about feature values is whether they should be expressed in exact terms or in terms of intervals or sets. For example, should a feature value of the blood pumping of hearts be the exact force by which it contracts or a force interval in which the force of the heart's contraction is located, or alternatively a set of exact forces or force intervals in which the force of the heart's contraction is located? This should be allowed to differ between feature types. In some cases, further life expectancy, or further expected number of offspring, varies with very small variations in feature values (given some construction of the organism fixed in all other respects). In such cases, fairly exact feature values are needed to describe the quantitative performance. In other cases there are ranges of exact values that give the same further life expectancy, or further expected number of offspring (given some construction of the organism fixed in all other respects). In such cases, the feature values should be expressed by intervals or sets.

The DPE-theory's allowance for intervals and sets of feature values makes it explicit that sometimes it is irrelevant for health and pathology what exact value within a certain interval or set is taken for a performance of a certain physiological function, given certain performances of other physiological functions. This means that the most well-integrated relational feature value for the performance of a physiological function need not always be an exact value, but a value within a certain interval or set. Hence, the allowance for intervals and sets of feature values plays a role in taking the importance of integration for health into account.

3.2.2 Feature Value Configurations

Since there may be several relevant feature types for a physiological function, we need to consider configurations of feature values. For example, if we evaluate the health status of a heart token, we probably need to consider not only its beating frequency, but also feature values for other relevant feature types, e.g. contraction force. This means that we need to look at configurations of feature values for a physiological function's relevant feature types. One can think of such a configuration as a list associating values to relevant feature types.

Remember that $RFT(F)$ denotes the set of relevant feature types for the physiological function F. We define a 'feature value configuration' as follows:

> A **feature value configuration** for a physiological function F is a function
>
> $fvc : RFT(F) \rightarrow \text{values}$
>
> such that for each feature type T in $RFT(F)$, $fvc(T)$ is a value for T.

That is, I use fvc to denote a feature value configuration and $fvc(T)$ to denote the feature value for the feature type T.

3.2.3 Specific Disposition Profiles

On the disposition profile efficiency theory, a trait token's health status depends on the efficiency of its dispositions for performing physiological functions. I will model such dispositions as "specific disposition profiles". Specific disposition profiles can be described as tables reporting, for each possible situation, how the trait token will respond with regard to one of its type's physiological functions; that is, in accordance with what feature value configuration (i.e. to which degree) the trait token will perform a physiological function. The idea is illustrated by the following table reporting a token heart's disposition for pumping blood (where each S is a situation, and each v and v' are values reporting on quantity of beating frequency and contraction force, respectively):

	S_1	S_2	…	S_n	…
Beating frequency	v_1	v_2	…	v_n	…
Contraction force	v'_1	v'_2	…	v'_n	…

We define a 'specific disposition profile' as follows:

A **specific disposition profile** for a physiological function F is a function

dp : situations → feature value configurations

such that $dp(S)$ is a feature value configuration for F for each situation S.

An important notion in this definition is that of a situation. By a "situation", I mean a possible state of the world. This understanding of situations is crucial for the DPE-theory for two reasons: As I will explain below, it is needed in order to solve Kingma's dilemma. Also, it is needed in order to account for the importance of integration for health.

We saw in the last chapter that Kingma's response to Hausman's answer to Kingma's dilemma does not succeed because Kingma does not consider states internal to the organism (e.g. states due to fertilization or breastfeeding or darkness) as part of the situation to which the trait token responds. If one includes such states in the situation, there is no problem to account for, for example, typical performances during pregnancy as healthy. It is, for example, inevitable for most women to have a lower activity of the immune system in the situation of having a fetus in the uterus, having certain hormone levels etc. Hence, pregnant women's specific disposition profiles are generally not different from non-pregnant women's.

Let us also revisit the case where Carol has ingested a poison that paralyzes her digestive system. I argued that Carol's digestive system should count as pathological only if its disposition for digesting food, after the ingestion of poison, is worse for survival or reproduction than what is required for health (for example by not digesting in any situation). One might perhaps think that Carol's digestive system cannot be accounted for as pathological in such a case, since it is typical for digestive systems to be disposed to not digest in any situation *after having poison ingested*. However, specific disposition profiles do not contain sequences of situations, but only specific situations. Hence, the only thing that matters for health are actual dispositions to perform physiological functions in different situations, not dispositions that will be acquired in some intermediate situation in a sequence of situations. Hence, it is irrelevant for health whether or not most digestive systems will acquire the same dispositions as Carol's digestive system in a situation where there is poison in the digestive system. As long as most digestive systems have not been in a situation where there is poison in the digestive system, a digestive system disposed to not digest in any situation will count as pathological.

Also in order to account for the importance of integration for health, the feature value configuration that a trait token's specific disposition profile takes must be considered relative to internal factors. Remember the example from Chapter 1, about the system regulating the body's overall use of energy, which includes the thyroid, the pituitary, and the hypothalamus. In order to evaluate the health status of a token pituitary, we cannot merely consider the amount of TSH that it secretes. Rather, we need to consider the amount of TSH it secretes in relation to the amount of THR secreted by the hypothalamus and in relation to the amount of T3 and T4 in the blood stream. The specific disposition profile of a token pituitary reports the amount of TSH secreted by that token for each combination of a certain amount of THR secreted by the hypothalamus and certain amounts of T3 and T4 in the blood stream (and also in combination with other internal or external factors). Hence, by considering specific disposition profiles rather than mere performances, the DPE-theory accounts for the importance of integration for health.

Since the term 'specific disposition profile' will be extensively used throughout the thesis, I will abbreviate it 'SDP'.

3.2.4 Complete Disposition Profiles

We have just seen how to represent a trait token's functioning with regard to a certain physiological function as a specific disposition profile. We may now use collections of SDPs to model the functioning of whole individuals. The idea is that we can represent the functioning of an individual by associating each of the individual's trait tokens' type's physiological functions with SDPs. I will call such a collection of SDPs a 'complete disposition profile'. Formally, I define a 'complete disposition profile' as follows:

> A **complete disposition profile** for a set P of physiological functions is a function
>
> $$DP : P \rightarrow \text{specific disposition profiles}$$
>
> such that $DP(F)$ is a specific disposition profile for the physiological function F for each F in P.

For illustration, consider a fictive organism having only a heart and a kidney. The complete disposition profile representing the functioning of this organism will include an SDP for each of the heart's physiological functions and an SDP for each of the kidney's physiological functions. The idea is illustrated by the

following table reporting the complete disposition profile of this individual (where each dp is an SDP). Here we assume that the only physiological function for hearts is to pump blood, and that the only physiological functions for kidneys is to filter blood and to release the hormone renin (affecting blood pressure).

Pump blood	dp_1
Filter blood	dp_2
Release renin	dp_3

I will use the abbreviation 'CDP' for 'complete disposition profile'.

In Section 3.1.2, I suggested that we measure survival chances by further life expectancy and reproduction chances by further expected number of offspring. I also discussed how to understand these measures for the purpose of the disposition profile efficiency theory. Since a CDP represents a whole organism, I will denote the further life expectancy of an organism with the CDP DP '$fle(DP)$', and I will denote the further expected number of offspring of an organism with the CDP DP '$feno(DP)$'. I will assume that further life expectancies are given by real numbers and that further expected number of offspring are given by natural numbers.

4 Reference Classes and Standards for Health

The previous chapter presented models of dispositions to perform physiological functions – specific disposition profiles (SDPs) and complete disposition profiles (CDPs). This chapter discusses the standard for healthy trait tokens – which CDPs or SDPs serve as a comparison standard in health status evaluations?

In accordance with earlier arguments for reference class-relativism, I will present a health standard which is relative to the reference class of the evaluated trait token's bearer (Section 4.1). I will also present an account of reference classes (which I presented some assumptions about already in Chapter 3) (Section 4.2.). We will see that the account of a reference class-relative standard and the account of reference classes are closely connected.

4.1 A Reference Class-Relative Standard for Health

Since I require of the account of health and pathology being developed that it is reference class-relative, I will define a health standard relative to reference classes. The standard for health discussed here should not be confused with a line separating health from pathology. The idea is that there is a standard for health, and that the line separating health from pathology is drawn somewhere in relation to this standard. One might have the idea that any deviation below the standard should be judged pathological, but that need not be the case. So, if a trait token is working at a level below the standard for health, it is not implied that the trait token is pathological. That depends on where, in comparison to the standard, the line separating health from pathology is drawn. I will discuss where this line should be drawn in the next chapter (Section 5.2.1).

I will first discuss an earlier suggestion of a reference class-relative standard by Hausman (2012) (Section 4.1.1), and then propose a reference class-relative standard for the disposition profile efficiency theory (Section 4.1.2).

4.1.1 Hausman's Standard

The disposition profile efficiency theory is not the first theory to use the idea of a reference class-relative standard for health. Since Boorse leaves the notion of efficiency largely undefined, is not clear whether the biostatistical theory uses a reference class-relative standard (in the sense I have in mind) to define 'health' and 'pathology'. Hausman (2012, 536), however, expresses an idea of a reference class-relative standard as part of his "functional efficiency theory". This is a theory which Hausman presents as an improved version of the BST (2012, 520). Hausman accounts for the standard in terms of what is attainable, or readily attainable, for many individuals of the reference class. Hausman writes:

> [w]hen there is a state of perfect part functioning, which is in some significant way different from other part behavior and moreover readily attainable in some common environments, then that state sets the standard of health, and any lesser functional efficiency counts as pathological (although possibly mildly so). (2012, 536)

An example of an optimal state, Hausman thinks, is "the possession of a complete set of teeth with no decay or other defects". This state is, according to Hausman, clearly more efficient than other possible states of the teeth, and it is "attainable by a nonnegligible number of human beings in some common environments" (2012, 536). When there is no single state of perfect part functioning as there is for the teeth, Hausman explains, the states that are readily attainable, and that have maximal efficiency in some benign, common, and relevant environments, count as healthy (2012, 536).

When I define a reference class-relative standard below, I will use Hausman's rough idea of the standard, but adapt it to the framework developed for the DPE-theory, and introduce some further precision.

4.1.2 Exemplary Complete Disposition Profiles

Defining a reference class-relative standard for health requires assumptions of how individuals of the reference class typically are, should be, or would best be, constructed. If there is no common construction for the reference class to relate to, one cannot say, for example, that a certain formation of teeth is best for some goal (e.g. survival). Then, whether a certain formation of teeth is optimal or not will depend, for instance, on the formation of the rest of the mouth, whether the organism may utilize solid food, etc. This means that,

when evaluating the health status of a trait token, we cannot simply adhere to an isolated standard for the dispositions for the trait type's physiological functions. For example, when evaluating the health status of a token set of teeth, we cannot just compare the dispositions of this set of teeth to the dispositions of some set of teeth taken in isolation. Rather, we need to compare it to the dispositions of a set of teeth that is part of an individual. Since an individual, according to the disposition profile efficiency theory, is modelled as a CDP (complete disposition profile), the standard will be accounted for as a CDP.

I will call the standard CDP 'exemplary'. To be more precise, I will distinguish between 'survival-exemplary' CDPs and 'reproduction-exemplary' CDPs. I will assume that, for each reference class, there is a survival-exemplary CDP and a reproduction-exemplary CDP. In Section 4.2, reference classes will be defined in a way that guarantees that there is, for each reference class, a survival-exemplary CDP and a reproduction-exemplary CDP. What does it mean, then, that a CDP is survival- or reproduction-exemplary? Roughly and informally, it means that the CDP is the best CDP to have for a member of the reference class in order to have as high further life expectancy, or, respectively, as high further expected number offspring, as possible, within reasonable frames.

I define a 'survival-exemplary CDP' as follows:

> A CDP DP is **survival-exemplary**, relative to a reference class R, if and only if
>
> out of the CDPs that are readily attainable for a significant share of the organisms in R in common environments for the organisms in R, DP gives the highest further life expectancy.

There are several parts in this definition that need further explanation or comment. I discussed how to understand "further life expectancy" in the last chapter (Section 3.1.2). However, it should be made clear, first, how to understand "readily attainable". Second, there is a question what counts as a significant share of organisms. Third, it should be clarified what "common environments" means.

The use of the term 'readily attainable' in the above definition is rather technical. By saying that a CDP is readily attainable for an individual, I mean that there is, among a set of relevant close nearby possible worlds, at least one world where the individual has the CDP. What limits the set of relevant possible worlds here? When discussing determinations of further life expectancy

and further expected number of offspring in Chapter 3 (Section 3.1.2), I concluded that special medical and social interventions should not be taken into consideration. This applies to determinations of survival-exemplary CDPs as well. This is because we do not want a standard for health that is dependent on specially directed medical and social interventions. So, the set of possible worlds that determine whether a CDP is readily attainable for an individual is a set of fairly nearby worlds, except for those worlds that contain special medical or social interventions directed towards the individual. What counts as fairly close here? This is of course a question of degree. The set should include worlds that differ in availability of rather controllable actions. For example, consider an individual who smokes, but for whom it had not been too difficult to choose not to start smoking. The set should include a world in which this individual is not a smoker. However, if the actual individual is strongly disposed to start smoking given that an opportunity arises, and common environments are such that there are a lot of opportunities for smoking, then it is questionable whether the set should include a world in which the individual does not smoke. Exactly at which distance to draw the line determining which worlds that are to be included I partly take to be a matter of convention.

Observe that this technical explication of 'readily attainable' means that if a CDP is readily attainable for an individual, this does not say that the CDP is possible for the individual to obtain in the future. Rather, it says that the individual could have had the CDP. To illustrate this in a concrete example, consider an individual who has smoked for quite some time. Because of her smoking, her lungs are scarred. Alveoli that have taken damage from smoking do not heal from that damage. Hence, the smoking cannot be undone. Yet, it may still be readily attainable, in the relevant sense, for this individual to have lungs not scarred by smoking.

Another question of degree is what counts as a significant share of the organisms in a reference class. This share should be rather large. Otherwise the standard for health would be unreasonably high. But exactly what counts as a significant share of the reference class members here, I take to be a matter of convention.

Like Hausman, I refer to common environments in the definition. And like Hausman, I take environments to be less fine-grained than situations.[22] The thought is that a certain environment may be instantiated by several different situations. One way to think about common environments, then, is as sets of

[22] Hausman (2012) distinguishes between "environments" and "circumstances". By "circumstances", Hausman seems to have something similar to my "situations" in mind.

situations that are sufficiently statistically likely. Since situations are very fine-grained, there is probably no situation that is literally likely. However, some situations should be more likely than others. If we consider some human reference class, it is for example more likely that a situation for some reference class member contains the facts of being in 18 degrees Celsius, being in a wind of 3 meters per second, and having a chewed sandwich in one's stomach, than that the situation contains the facts of being in 18 degrees Celsius, being in a wind of 3 meters per second, and having a certain unusual poison in one's stomach.

Similar to the previous points about the definition, at which (low) probability level a situation becomes likely, I take to be a matter of convention.

In the above clarification of the survival-exemplary standard, conventions are used to set the exact boundaries determining survival-exemplary CDPs. In the further development of the DPE-theory, we will see that conventions are used to set exact boundaries determining some other components of the theory as well. The point of referring to conventions is that it does not matter for the purpose of the DPE-theory whether one particular boundary or another is chosen, as long as both boundaries fulfill certain requirements, or are located within certain limits. Irrespectively of which exact boundary is chosen, the theory captures what it is supposed to capture. Someone might perhaps worry that conventionally set boundaries, by being socially negotiated, allow for values to sneak into the theory. One might, for example, worry that certain ideas about which conditions that should be disvalued, or certain ideas about health care allocation, guide the choice of boundaries. However, although such ideas could motivate choosing a certain convention, this does not make the theory itself value-involving. In order to use the DPE-theory as a basis for disvaluation, or for health care allocations, it must be combined with some values. The DPE-theory does not by itself imply anything about disvalue of certain conditions, or about health care allocation.

Similarly to how we defined a 'survival-exemplary CDP', we may define a 'reproduction-exemplary CDP':

> A CDP DP is **reproduction-exemplary**, relative to a reference class R, if and only if
>
> out of the CDPs that are readily attainable for a significant share of the organisms in R in common environments for the organisms in R, DP gives the highest further expected number of offspring.

The definitions above allow that a reference class has several survival-exemplary CDPs and several reproduction-exemplary CDPs. I will use '$\boldsymbol{S\text{-}Ex}(R)$' to denote the set of all survival-exemplary CDPs of the reference class R. I will use '$\boldsymbol{R\text{-}Ex}(R)$' to denote the set of all reproduction-exemplary CDPs of the reference class R.

Should we expect reference classes to have many different survival-exemplary and reproduction-exemplary CDPs? Here, we should first note that the number of survival- and reproduction-exemplary CDPs are limited in an important way by a theoretical choice discussed in Chapter 3 (Section 3.2.1), namely the choice to allow for expressing feature values in terms of intervals or sets. Sometimes it does not matter for further life expectancy, or further expected number of offspring, by what exact quantity a physiological function is performed. Let us assume that it does not matter at all for further life expectancy whether a human heart is disposed to beat with 50 or 52 beats per minute (or a beating frequency in between) when the individual is resting. On the DPE-theory, this span of exact beating frequencies should then be clustered together, so that the survival-exemplary CDP's SDP (specific disposition profile) for pumping blood takes the value 50-52 beats per minute for the feature type associated with beating frequency in the situation of resting. Hence, we will not have several survival-exemplary CDPs that differ in their exact beating frequencies within the interval 50-52.

There may, however, still be reference classes with several survival- or reproduction-exemplary CDPs, at least in theory. To see this, consider the following example. Say that, in a reference class, there are two trait types which each has one physiological function. For the first physiological function there is one relevant feature type, call it '$FT1$'. For the second physiological function there is also one relevant feature type, call it '$FT2$'. Regarding $FT1$, there are two alternative values that are readily attainable for a significant share of the reference class members in common environments of the reference class; call them 'A' and 'B'. Regarding $FT2$, there are also two alternative values that are readily attainable for a significant share of the reference class members in common environments of the reference class, call them 'C' and 'D'. From this we have four possible combinations of feature values: $A + C$; $A + D$; $B + C$; and $B + D$. Assume that it also holds for all four combinations that they are readily attainable for a significant share of the organisms in the reference class in common environments of the reference class. Let us assume that, with regard to further life expectancy, the combinations $A + C$ and $B + D$ are most beneficial, and neither is more beneficial than the other. This means that A and B, and respectively C and D, cannot be thought of as part of the same relevant

feature value interval or feature value set. In this example, there are two survival-exemplary CDPs for the reference class, one with the combination $A + C$ and one with the combination $B + D$.

I do not have any real-life example of feature values like A, B, C, and D. But if there are such real examples, they will presumable not amount to very different survival- or reproduction-exemplary CDPs of a reference class. Given that organisms that share physiological functions are rather similar, the different survival- or reproduction-exemplary CDPs will probably be rather similar. How to handle the potential existence of several survival- or reproduction-exemplary CDPs will be discussed in the next chapter (Section 5.2.2).

4.2 Reference Classes

I have now accounted for a reference class-relative standard for survival-health and a reference class-relative standard for reproduction-health. We will see that these standards put some requirements on reference classes. In Chapter 3 (Section 3.1.4), I listed some things I assume about reference classes in the disposition profile efficiency theory. Let us now look in more detail on how to account for reference classes in a way that may serve as a basis for health status evaluations. An important point about the account that I will present is that it is not meant as a general account of reference classes in biology. The purpose of it is to define a relevant comparison class for the evaluation of the health status of a trait token in an organism.

In contrast to the biostatistical theory, the account does not primarily aim to describe how reference classes tend to look like in physiology and pathology. Rather, the account offers an *explanation* of why comparison classes of organisms tend to look like they do in health status evaluations, for example why different species, different sexes, and organisms in different stages of life tend to belong to different comparison classes.

The approach of the account that I will propose is holistic. The basic idea is to search for a best overall fitting division of organisms into reference classes, based on five ideas. These ideas are motivated, first, by basic interests in physiology. Second, they are motivated by theoretical requirements imposed by the role reference classes are supposed to play in the disposition profile efficiency theory. The reference to basic interests in physiology explain why the reference classes which the account results in are interesting to consider. The reference to theoretical requirements is called for, since if the account of reference classes does not fulfill these requirements, then the account will not

be meaningful for a theory of health and pathology anyway. Is it problematic that the role reference classes are supposed to play in health status evaluations affect the individuation of reference classes? I think it is not. It would be problematic if the account aimed for reference classes functioning as general divisions of organisms in biology. But this is, as mentioned above, not the case here. They need only be relevant comparison classes for health status evaluations.

So, in response to the worry discussed in Chapter 2 (Section 2.3.1), about values motivating accounts of reference classes, I think the motivation of the account should be fine. In any case, a clear, transparent motivation should be considered beneficial since it makes the theory open to critical examinations, both regarding scientific interest and moral aspects.

The first idea concerns what may be described as the main interest in physiology and pathology, namely structure and functionality. The division of organisms should be guided by grouping together organisms that are similar in structure and functionality directed towards survival and reproduction.[23] In order to preserve as great variation in physiological functionality as possible, the division should be guided by a goal of assigning physiological functions to as many body part tokens as possible. Hence, alternative ways of functioning within a group of organisms is a reason to consider the different organisms as belonging to different reference classes.

The second idea concerns reference class-relativism: In order for a set of organisms to count as a reference class, it must be possible to determine a standard for survival-health and a standard for reproduction-health for this set of organisms. Otherwise, the reference classes that the account provides cannot serve as basis for health status evaluations (given that health and pathology, as I have argued, are reference class-relative properties).

The third idea is related to the idea about reference class-relativism, and concerns stability. The reference class-relative standards for survival-health and reproduction-health should be rather stable, i.e. they should not be too sensitive to temporal fluctuations. With too unstable reference classes it becomes questionable whether health status determinations are meaningful to make.

[23] This account of reference classes is formulated based on the assumption that the relevant goals for physiological functions are only survival and reproduction, and that survival chances and reproduction chances are best measured in terms of further life expectancy and, respectively, further expected number of offspring. If these goals or measures are changed, the account needs to be adjusted accordingly.

When arguing against the biostatistical theory, some (e.g. Guerrero 2010) have argued that if something rather suddenly happens to most members of a reference class, so that what is typical in the reference class changes for the worse (say, for example, that a new virus makes every human blind), then this state should count as pathological. According to Guerrero (2010), this becomes problematic for theories like the BST, where the health status of a trait token is relative to the functioning of the other tokens of that trait type in the reference class. If what is normal changes for the worse, then the health status of a particular trait token may change although the trait token does not change at all.

I agree with the intuition that if something changes drastically for the worse for all members of a reference class overnight, or even over a few generations, then the new common, worse, state should still count as pathological. However, if the new worse state continues to be the best readily attainable state for quite some time (comprising many generations), then it should reasonably no longer be considered pathological.[24]

The fourth idea is also related to the idea about reference class-relativism, and concerns age: the general standards for survival-health and reproduction-health should be age-relative. As we saw in Section 4.1, a survival-exemplary CDP (complete disposition profile) of a reference class is a CDP that gives the highest further life expectancy, among those CDPs that are readily attainable for a significant number of organisms in the reference class. However, as noted already in Chapter 3 (Section 3.1.2), the further life expectancy for an individual will generally vary depending on its age. We saw that this is because the further life expectancy of an organism generally decreases during its life. For example, an 80-year-old human individual typically has a much lower further life expectancy than a 10-year-old human individual. However, if we consider a perfectly normal organism throughout its life, the judgement should not be that its different trait tokens get more and more pathological the older the organism gets, merely because they get associated with lower and lower further life expectancies. However, this will be the judgement if we do not relativize survival-exemplary CDPs to age. Then we will have to say that the standards for survival-health for human reference classes are those of babies or children, and that most humans older than babies or children comprise a lot

[24] For further discussion along the same lines, see Boorse's (2014, 717-716) response to Guerrero's (2010) objection.

of serious pathology.[25] A similar story could be told for reproduction-exemplary CDPs. If reference classes were not individuated by age, this would for instance mean that the reproductive organs in most women older than 60 are reproduction-pathological. Relating to basic interests in physiology, normal changes in structure and functionality during organisms' life cycles are of physiological interest.

The fifth idea is related to the idea about structure and functionality. It concerns evolutionary aspects. With time, environments change and, often in response to such changes, structure and functionality change. For example, consider the wisdom teeth in the human species. For human ancestors, wisdom teeth served a physiological function which they do not today. The diet in common environments of the human ancestors required more efficient chewing than does the diet in common environments of more modern humans. Today, missing wisdom teeth does not negatively affect survival. For a human ancestor, on the other hand, it did. The differences need not, as in the example, be between having a certain body part with a certain function and not having that body part with that function. Rather, the difference may be merely between having a certain body part disposed to carry out a certain function with a certain intensity and having a body part, that is of the same type but disposed to carry out the function with a different intensity.

Because of structure and functionality changes involved in the evolutionary process, reference classes should be limited by time boundaries. The division should not put organisms separated in time by more than a certain number of generations in the same reference class. The number of generations referred to here should be small enough to capture stable evolutionary variations. Exactly what this amounts to is, however, a question of convention. A similar evolutionary aspect concerns evolutionary separated groups of organisms. The division should not put organisms belonging to evolutionarily separated groups of organisms, i.e. organisms that have not had a common ancestor for the certain – conventionally determined – number of generations, in the same reference class.

The account of reference classes that I will suggest makes assignments of reference classes to individuals relative to time, in two senses. First, reference

[25] In Chapter 3 (Section 3.1.2), I discussed whether we should measure survival chances by further life expectancy or simply by life expectancy. We may note that even if we measured survival chances by life expectancy, we would have to individuate reference classes by age. This is because, as mentioned in Chapter 3, the life expectancy generally increases during an organism's life. If we did not individuate reference classes by age, we would get the result that organisms generally get healthier the older they get, merely by already having achieved a certain age.

classes are sets of individuals viewed at specific time intervals. This implies that a particular organism does not just belong to a certain reference class; rather, it belongs to a certain reference class during a certain time interval. During the lifetime of an organism, the organism will have belonged to many different reference classes. The reason for this time-relativity is the above-mentioned idea that reference classes should be individuated by age. An individuation of reference classes by age requires that organisms are viewed at specific time intervals, and that an organism can belong to different reference classes at different time intervals.

What is the length of the time intervals referred to here? In order to capture well variations in further life expectancy and further expected number of offspring, the time intervals should be rather short. However, in order to not go into unnecessary detail about such variations, the intervals must not be extremely, or infinitely, small.

The second sense in which the account is time-relative is that it individuates reference classes relative to time intervals. Depending on the time interval chosen, different individuals will be included in the division. The relevant time interval to consider in a health status evaluation of a trait token is that in which the health status is taken to apply to the trait token. The idea of the account is to take into consideration all organisms that live at this time interval, plus all organisms that have lived not more than a certain number of generations back in time from this time interval, and find a best overall fitting division of them into reference classes. (In accordance with the first sort of time-relativity, the division includes versions of these individuals in all stages of their lives.) The reason for this time-relativity is connected, first, to the ideas about reference class-relativism and stability, and, second, to the idea that reference classes should be limited by time. By including more individuals than those presently alive, we get a larger statistical basis for determinations of survival- and reproduction-exemplary CDPs. Also, in order to make reference classes stable, so that survival- and reproduction-exemplary CDPs do not become too sensitive to temporal fluctuations, the account takes into consideration more organisms than those that live at the relevant time interval. By taking into consideration several generations of organisms, the account handles sudden changes, for example a suddenly widespread virus causing blindness. As long as the virus has not been around so long that it is no longer readily attainable for a significant share of the organisms in the reference class to see in day light, being blind will still count as pathological. However, if the virus continues to be around for a long time, blindness will cease to be pathological.

The reason why the account of reference classes does not take into consideration all organisms that have ever lived has to do with the evolutionary aspects discussed above. The division should not include in the same reference class organisms separated in time by more than a certain number of generations. Hence, the account will balance between, on the one hand, including as many individuals as possible and not being too sensitive to temporal fluctuations and, on the other hand, accounting for functional differences arising along the history of evolution.

Let us now consider how the individuals taken into consideration are divided into reference classes. Below, I will present six principles according to which the division is done.

First, we have two principles connected to the idea of reference class-relativism. These principles are needed to ensure that, for each reference class, there is a survival-exemplary CDP and a reproduction-exemplary CDP. The first principle is the following:

> **Principle 1.** Each reference class must contain a sufficiently large number of individuals.

This principle is needed for statistical reasons. In order for a set of individuals to count as a reference class, the set must contain enough individuals to serve as a statistical basis for meaningful judgements about typicality of physiological functions and trait types. If a reference class consisted of only one individual, it would not make sense to make claims about what is typical in that reference class. Even if there are ten individuals, there is not much of a statistical basis for meaningful judgements about what is typical. Although I will not make it more precise here what counts as a sufficiently large number, we can at least conclude that a reference class must consist of more than one or ten organisms. In order to get a more precise number, we could consult statistical theory. We may, however, note that what is a large enough number is somewhat vague. Exactly where to set the bar will be partly a matter of convention.

The second principle concerns the determination of trait types and physiological functions in a reference class:

> **Principle 2a.** For each reference class, there is exactly one set of organism part types with associated function types that together are readily attainable for a significant share of the reference class members in common environments of the reference class.

Principle 2b. Each of these functions contributes to the further life expectancy of a survival-exemplary CDP of the reference class, or to the further expected number of offspring of a reproduction-exemplary CDP of the reference class.

The types of organism parts referred to in Principle 2a-b are trait types, and the associated function types are physiological functions. It may be noted here that my use of 'trait' does not correspond exactly to how this term is used in biology. In biology, a 'trait' may refer to a characteristic of an organism in the sense of a behavior. On the DPE-theory, a trait is always a body part. With regard to behaviors, the brain (or some specific area of the brain) is usually responsible. It is the brain, then, rather than the behavior, that is the trait, and it is a function of the brain to produce the behavior.

Remembering Boorse's terminology, Principle 2a and 2b concern the question of what strength is required for physiological functions. What counts as a "significant share" of organisms is, however, vague. Although I will not suggest a precise function determining the number of instances needed for there to be a significant share, we may note that the intended share should balance between not requiring that too many have the physiological function readily attainable (since that, for example, would exclude organisms with congenital or acquired deficiencies) and not allowing that too few have the physiological function readily attainable (since it would then not be typical to have that physiological function readily attainable).

A further question one may have regarding Principles 2a and 2b is what it means that a function contributes to the further life expectancy, or, respectively, to the further expected number of offspring of an organism. We saw that Boorse defines 'physiological functions' with reference to causal contribution to survival and reproduction. However, he never defines what it means to causally contribute to these goals. That a function contributes to the further life expectancy or, respectively, to the further expected number of offspring of an individual, I take to mean that the function, given the construction of the organism in other respects, makes a positive difference for further life expectancy, or, respectively, for further expected number of offspring. That is, the contribution may be measured by considering the difference in further life expectancy, or, respectively, in further expected number of offspring, between the organism in question as it is and the organism as it would be if it did not have the token function in question.

With regard to Principles 2a-b, we may also note that the DPE-theory handles polymorphisms differently than does the BST. We saw in Chapter 2 that Boorse allows for alternative normal ways of functioning within a reference class by including polymorphisms disjunctively in the species ideal. According to Principles 2a-b, in contrast, each reference class will have exactly one set of physiological functions. If there are two alternative functions that are both common among a group of otherwise similar individuals, then these individuals do not all belong to the same reference class. However, it is not clear that all trait variations classified as polymorphisms in biology are functionally (i.e. qualitatively) different from each other. For example, consider a typical example of polymorphic variants, blue and brown irises. Blue irises differ from brown irises by letting in more light to the eye. This is, however, only a quantitative difference. The physiological function of the respective irises is the same – to hinder light from entering the eye. Hence, the DPE-theory does not individuate reference classes by differences in eye color. But when there are polymorphic traits that have qualitatively different functions, they individuate reference classes.

The third principle is motivated by the idea that survival-health and reproduction-health are relative to age:

Principle 3. Each reference class only contains individuals of the same age.

Principle 3, in connection with Principle 1, may leave some individuals without a reference class. If the number of organisms of a certain age, with certain organism part types and associated function types, is too small, then there is no reference class to which these individuals belong. Is this implication problematic? It means that we cannot evaluate the health status of these organisms' parts. Although this implication may be regarded as unintuitive, I do not consider it too problematic. Note, first, that the number of organisms for which this will be the case are very few, at least in relation to the number of organisms that will belong to some reference class. Since the division takes into consideration organisms in all stages of their lives, there will for most ages be at least a minimal number of organisms. It is only the very extreme individuals, that are older than at least almost everyone else, or are members of a species that has been under severe decimation for a timeframe of many generations, that will be left without a reference class. And although it would in a sense be nice to be able to evaluate the health status of different parts of these organisms, I think it is also a rather natural conclusion – given the acceptance of a statistical measure – that the basis for health status evaluations is too thin

for it to be meaningful to judge parts of these individuals survival- or reproduction-healthy or -pathological. What one could perhaps say is that a body part of an extremely old organism, say of a 130-year-old human individual, is survival-healthy or survival-pathological if viewed as if she was 100 years, or that an organism of a species under extinction is survival-healthy or survival-pathological if viewed as contemporary with individuals of the species at the time when the species was not near extinction. However, the lack of a sufficient statistical basis will make it impossible to determine unconditional health statuses of these organisms.

The fourth principle is motivated by the second evolutionary aspect mentioned above:

> **Principle 4.** No reference class includes individuals belonging to evolutionarily separated groups, i.e. individuals that do not have a close enough common ancestor.

The number of generations distinguishing a close enough common ancestor from one that is not close enough, I take to be the same as the number of generations determining which organisms that should be taken into consideration in the division of organisms into reference classes.

The fifth principle is motivated by the ideas about structure and functionality. It is needed to preserve as great variation in functionality as possible:

> **Principle 5.** Given Principles 1-4, the division of reference classes maximizes the total number of trait types among the individuals taken into consideration in the division.

With this principle, the division should result in as many different reference classes as possible, as long as these classes tracks alternative ways of functioning and fulfil Principles 1-4. Importantly here, two sets of functions, A and B, count as "alternative ways of functioning" if and only if A contains at least one function that B does not and B contains at least one function that A does not.

The sixth principle is technical in character. The purpose of it is to avoid further division of reference classes than required by Principles 1-5.

> **Principle 6.** The division of individuals generates no more reference classes than is needed to fulfill Principles 1-5.

A minor reason for this principle is to keep the set of reference classes as small as possible, when this does not have any cost for the theory (besides the complication of an additional requirement). A more important reason for the principle is that reference classes should not be further individuated based on other considerations than those already mentioned. For example, if the division of organisms into reference classes was not limited to Principles 1-5, it would be possible to have a division such that amputating a foot makes one end up in a reference class where everyone has only one foot. This would mean that the lack of a foot would not make one's leg survival-pathological, which is obviously a strange result.

We are now in a position to define a 'reference class', relative to an organism. I will denote an organism O at a certain time interval t 'O-at-t':

> An organism O **belongs to the reference class** R, at the time interval t, if and only if R is assigned to O-at-t in accordance with the division of all organisms that live at t or have lived not more than a certain number of generations back in time from t in accordance with Principles 1-6.

Since there may be several divisions of reference classes that all fulfill Principles 1-6, an organism may, at a specific time interval, belong to several reference classes. I will use '$\boldsymbol{R}(O,t)$' to denote the set of all reference classes to which the organism O belongs at the time interval t.

Should we expect that organisms typically belong to several different reference class at specific time intervals? Given the biological world, I assume that cases where an organism belongs to more than one reference class should be rare. But, more importantly, in such potential cases, the different reference classes should be rather similar. For example, if a human male at the age of 35 belongs to two different reference classes, these will not be as different as one comprising only humans and the other comprising mainly chimpanzees. Rather, the difference will probably concern exactly which set of human males at the age of 35 that are included in the reference class. In the next chapter (Section 5.2.2), I will discuss how to handle the potential case where the individual, whose trait token's health status we evaluate, belongs to more than one reference class.

5 Efficiency, Health, and Pathology

In the previous chapter, I accounted for a reference class-relative standard for health. This chapter concerns how to use this standard in health status evaluations. I have said that when we evaluate a trait token's survival-health status, we compare the trait token under evaluation to the standard. In this chapter I will explain how this is done.

I will call the model for comparing the trait token under evaluation with the standard an efficiency measure. The first part of this chapter (Section 5.1) defines and discusses this efficiency measure. The second part of the chapter explains how to use the measure in health status evaluations (Section 5.2). Roughly, the idea is that if the efficiency is high enough, it indicates that the trait token is healthy, and if the efficiency is not high enough, then the trait token is pathological.

5.1 Efficiency

In Chapter 4 (Section 4.1.2), I distinguished between survival-exemplary CDPs (complete disposition profiles) and reproduction-exemplary CDPs. Accordingly, we need to distinguish between survival-efficiency and reproduction-efficiency. For health and pathology with regard to survival, survival-efficiency is the relevant measure, and for health and pathology with regard to reproduction, reproduction-efficiency is the relevant measure. To keep it shorter and not too messy, I will here focus on survival-efficiency. The thought is that we can account for reproduction-efficiency analogously. Below, I will first discuss efficiency as a feature of dispositions (Section 5.1.1). Then, I will present Hausman's suggested definition of 'efficiency', which accounts for efficiency as a feature of dispositions (Section 5.1.2). Hausman's account is, however, problematic. Lastly, I will suggest how to define 'efficiency' in a way that avoids the problems of Hausman's definition (Sections 5.1.3-5.1.4).

5.1.1 Efficiencies of Dispositions

When we evaluate a trait token's survival-health status, we compare the trait token under evaluation to the standard. More precisely, we compare the dispositions of the trait token under evaluation to the corresponding dispositions of the standard. Since we are interested in health and pathology with regard to survival, the respect in which we compare them is further life expectancy. If we were instead interested in health and pathology with regard to reproduction, we would make the comparison with regard to further expected number of offspring.

Since a trait type may have several physiological functions, and a trait token's dispositions for performing these functions may be better or worse with regard to survival, we have to consider them one at a time. That is, when evaluating the health status of a trait token, we must determine the efficiency of the disposition for each of the physiological functions of the token's trait type. Dispositions for specific physiological functions are, as we saw in Chapter 3, modelled by specific disposition profiles – SDPs. Hence, we will measure efficiency as a property of SDPs.

To look at the different physiological functions separately when evaluating the trait token's health status is in line with the biostatistical theory. However, an important difference from the BST is that the disposition profile efficiency theory ascribes efficiencies to *dispositions* for performing physiological functions, rather than to performances of physiological functions. The reason for defining 'efficiency' as a feature of dispositions, rather than a feature of performances, is that it makes more sense to measure an organism's further life expectancy based on information about how its different trait tokens are disposed to perform in different possible situations, than merely on information about how its trait tokens are disposed to perform in one particular situation.

Someone who accounts for efficiency as a property of dispositions, or of capacities, is Hausman (2012). Let us next consider his definition of 'efficiency'.

5.1.2 Hausman's Account of Efficiency

As part of his "functional efficiency theory", Hausman suggests the following definition of 'efficiency':

> On the assumption that the other parts of the organism or systems whose activities do not depend on [the set of capacities] C or [the set of capacities] C′ are functioning adequately in a relevant environment,

> The functional efficiency of C is greater than that of C′ in some system S with respect to some goal G if and only if C makes it more likely that S achieves G than does C′. (2012, 534)

Compared to Boorse's notion of efficiency, this definition is different, first, in that Hausman defines 'efficiency' as a feature of sets of capacities, rather than of performances. Second, Hausman does not, as Boorse, limit the system to an organism, or the goals to survival and reproduction. However, it is clear that he takes organisms to be relevant systems, and survival and reproduction to be relevant goals for health status evaluations (see Hausman 2012, 521, 535). Third, the definition is more precise by stating that efficiency is about the likelihood of achieving a goal of a system.

Although Hausman's definition has advantages over Boorse's, it is, I will argue, not suitable to apply in health status evaluations. The reason for this is twofold. First, it is questionable whether measuring efficiency in terms of the likelihood to achieve a certain goal always makes sense in health status evaluations. Consider the goal of survival. Just "achieving survival" is a too unspecific goal for meaningful efficiency comparisons. To be more specific, one could consider the likelihood to survive for a particular time interval. When evaluating treatments for cancer, one typically considers progression-free survival of certain time intervals (e.g. one year). However, health status evaluations do not only require considerations of the likelihood to survive for a certain interval of time, but considerations of likelihoods to survive for all different time intervals. To see this by an example, consider the individuals Astrid and Beatrice, both 20 years old. The capacities of Astrid's heart give higher chances of surviving for one year than the capacities of Beatrice's heart. But the capacities of Beatrice's heart give higher chances of surviving for 20 years than the capacities of Astrid's heart. Here we may conclude that the capacities of Astrid's heart are more efficient than those of Beatrice's heart, relative the goal of surviving for another year, but that the capacities of Beatrice's heart are more efficient than those of Astrid's heart, relative the goal of surviving for another 20 years. But it is unclear whether we should say that Astrid's or Beatrice's heart is healthier.

The second reason why Hausman's definition is not suitable to apply in health status evaluations is that it only defines 'efficiency' comparatively. Hausman provides us with an ordinal scale for measuring efficiencies, which allows us to rank the efficiencies of different sets of capacities. This makes it

possible to draw a line distinguishing health from pathology based on statistical normality. One could for example say that the 2,5 percentage of the tokens of a trait type in the reference class that have the lowest efficiencies in the ranking are pathological. However, this is precisely the sort of line-drawing that Hausman argues against. As we saw in Chapter 2 (Section 2.3.4), the biostatistical theory falls prey to the problem of common diseases because it defines 'pathology' based on the prevalence of efficiencies in a reference class. Hausman (2012) recognizes this and argues that the line distinguishing health from pathology is therefore not statistical. He writes:

> Nevertheless, the functional efficiency theory, in contrast to the BST, denies that prevalence defines whether a level of functional efficiency is healthy or pathological. Although what is prevalent is typically healthy, the problems of common diseases and healthy populations show that this coincidence is not perfect. (Hausman 2012, 536)[26]

Yet, with only an ordinal scale of efficiencies, it is doubtful whether Hausman can define a general line separating health from pathology that is not based on statistical normality of efficiencies. In order to do that, it seems that we need to define 'efficiency' using an interval or ratio scale. In Hausman's defense, it should be added that Hausman's primary goal is not to define an exact line distinguishing health from pathology. Rather, he expresses doubts towards the usefulness of such a line:

> Once physiologists have figured out how to assign levels of efficiency to the capacities of some part or subsystem within a reference class in a relevant environment, it does not matter much theoretically whether they classify any particular level of functional efficiency within some system or within the organism as a whole as pathological or healthy. (Hausman 2012, 534)

However, even if Hausman is right and it is not that important to distinguish clearly between health and pathology, we may still want to be able to make

[26] The problem of healthy populations is formulated by Schwartz (2007) as a problem very similar to the problem of common diseases. The problem is that if the line distinguishing health from pathology is defined in terms of prevalence in a reference class, then it is – unintuitively – impossible for all members of a reference class to be healthy in a certain respect. The reason why I do not discuss this problem, but only the problem of common diseases is, first, that there are more clear real examples of common diseases than of "healthy populations", and, second, that the problem of common diseases has gained much more interest within the debate than the problem of healthy populations.

more comparisons between different efficiencies than Hausman's account allows for. On Hausman's account we may conclude about three trait tokens of the same type that one works best, another second best, and the third worst. However, we cannot, for example, evaluate whether the second best and the worst differ much or only little in how much worse they function in comparison to the best functioning token. Another limitation of Hausman's account is that we cannot, at least not in the relevant sense, compare trait tokens with different physiological functions. It is not clear whether C and C′ in Hausman's definition of 'efficiency' must be capacities of the same type, or whether they may differ in what types of capacities they are. If they must be of the same type, it follows that we are not able to compare trait tokens of different types (e.g. a token heart and a token liver) with regard to their capacities' efficiencies. If C and C′ may represent capacities of different types, then we can compare trait tokens of different types with regard to their capacities' efficiencies. However, we are not able to compare them in the relevant sense. We may conclude, for example, that a specific heart's capacity to pump blood has a higher efficiency then a certain muscle's capacity to abduct the thumb, since the heart's capacity to pump blood makes it more likely that the individual achieves the goal of becoming 80 years old than does the muscle's capacity to abduct the thumb. But this is not the sort of comparisons we are interested in here. We do not want to say that a normally functioning muscle in the hand works worse than a normally functioning heart just because the heart plays a more central role in the survival of a human than does the muscle in the hand. What we are interested in is rather a comparison between, on the one hand, how the specific heart's capacity to pump blood relates to a standard for pumping blood, and, on the other hand, how the specific muscle's capacity to abduct the thumb relates to a standard for abducting the thumb. This sort of comparison cannot be made on Hausman's account.

We may also note that Hausman's definition explicitly makes comparative efficiency relative to a system, and exactly one system. When determining the comparative efficiency of two sets of capacities, these sets of capacities must be seen as alternative sets of the same specific system (e.g. as alternative capacities of a specific individual). This is, as I will discuss more in the next section, a good idea which is in line with reference class-relativism. However, Hausman does not explain how to compare two sets of capacities as alternative sets in one and the same system. And in his concrete examples, he does not really incorporate the idea. It is, for instance, not clear in the following example how Jill's heart and Joan's heart are compared as alternative hearts of one and the same individual:

> […] systems or organisms with higher levels of functional efficiency are, with regard to the particular part or process, healthier, and systems or organisms with lower levels of functional efficiency are less healthy. For example, if Joan's heart enables her to walk up stairs easily, while Jill's does not permit her to walk across the room without stopping to catch her breath, Joan's heart is healthier. (Hausman 2012, 534)

In this example, it sounds as if what is compared is, on the one hand, Joan's heart as part of Joan and, on the other hand, Jill's heart as part of Jill. But if we compared Joan's and Jill's respective hearts as alternative parts of one and the same system, we would have to view them as alternative hearts of Joan, or as alternative hearts of Jill, or as alternative parts of some third individual.

In the following sections, I will suggest a definition of 'efficiency' which serves health status evaluations better than Hausman's. In Section 5.1.3, I will suggest a way to view different SDPs (specific disposition profiles) for a certain physiological function as alternative parts of one and the same system of SDPs (i.e. of one and the same CDP (complete disposition profile)). In Section 5.1.4, I will define 'survival-efficiency' based on a more fruitful measure than the likelihood of achieving a goal G, namely further life expectancy (which was suggested already in Chapter 3 (Section 3.1.2)). This definition of 'survival-efficiency' will also let us order efficiencies not only on an ordinal scale, but on a ratio scale.

5.1.3 Comparing Dispositions

I have indicated that when evaluating the survival-health status of a trait token, we compare its SDPs (specific disposition profiles) to the corresponding SDPs of the survival-exemplary CDP (complete disposition profile) of the trait token bearer's reference class. I have also indicated that the respect in which we compare them is further life expectancy. Let us start to consider how this comparison is carried out.

Since further life expectancy and further expected number of offspring are features of CDPs rather than single SDPs, we cannot compare two SDPs taken in isolation. Rather, we must compare them as parts of CDPs. This also seems plausible. As I pointed out in Chapter 4 (Section 4.1.2), whether a trait token's disposition for a physiological function works well with regard to survival depends on the system of dispositions that it is part of. Which two CDPs, then, are we to compare to each other? The CDP that serves as comparison is, naturally, the whole survival-exemplary CDP. But in the context of which CDP should we view the SDP that we evaluate?

We should not view the SDP that we evaluate in its actual context, i.e. the CDP of the bearer of the trait token under evaluation. One reason for this is that if we compare the CDP to which the trait token under evaluation belongs to the survival-exemplary CDP, then we cannot isolate the effect on further life expectancy caused by the SDP under evaluation. If the further life expectancy of the CDP to which the trait token under evaluation belongs is lower than the further life expectancy of the survival-exemplary CDP, this does not imply that the SDP that we want to evaluate is not working as well as the corresponding SDP in the survival-exemplary CDP. The lower further life expectancy could be caused by other SDPs than the one under evaluation. Say, for example, that we evaluate Charlie's heart's SDP for pumping blood. Assume that Charlie's heart's SDP for pumping blood is exactly like that of the SDP for pumping blood in the survival-exemplary CDP. Now, assume also that Charlie's lungs are not working as well as in an individual with a survival-exemplary CDP. Here, we do not want to say that Charlie's heart's SDP for pumping blood has a reduced survival-efficiency merely because the lungs make Charlie's further life expectancy shorter than that of the survival-exemplary CDP.

Another reason why we should not view the SDP that we evaluate in its actual context has to do with the desideratum of reference class-relativism. As I argued in Chapter 1 (Section 1.1.6), I think that if two distinct heart tokens of two different individuals of the same reference class are exactly similar, then the heart tokens' health statuses should be the same, irrespective of the construction of these two individuals otherwise.

Instead of viewing the SDP under evaluation as part of the CDP that it actually is part of, we should view it as an alternative part of the survival-exemplary CDP. Hence, when we evaluate the survival-efficiency of an SDP we compare the further life expectancies of two CDPs: on the one hand, the reference class-relative standard for health, i.e. the survival-exemplary CDP, and on the other hand, a CDP resulting from substituting the SDP under evaluation for the corresponding SDP in the survival-exemplary CDP.

I will call the CDP that we get by changing a CDP with regard to a certain SDP 'a CDP manipulated with an SDP'. Let us define this formally. Given an SDP dp, I denote the physiological function that dp is an SDP for '$f(dp)$'. We then define 'the manipulation of a CDP with an SDP' as follows:

> The **manipulation** of a CDP DP with an SDP dp is the CDP DP' which is exactly like DP except that $DP'\big(f(dp)\big) = dp$.

I will denote the manipulation of the CDP DP with the SDP dp '$M(DP, dp)$'. When talking about "a CDP DP manipulated with an SDP dp" I have the same thing in mind.

To illustrate the idea, say that we are interested in evaluating the survival-efficiency of Dragan's cochlea's SDP for transforming sound waves into neural signals. This means that we have to manipulate the survival-exemplary CDP of Dragan's reference class with Dragan's cochlea's SDP for transforming sound waves into neural signals. To do this, we locate the SDP for transforming sound waves into neural signals in the survival-exemplary CDP. We substitute this SDP with Dragan's cochlea's SDP for this function. So, if Dragan's cochlea is not able to transform any soundwaves to neural signals (i.e. he is deaf), then the manipulated survival-exemplary CDP will also be unable to transform soundwaves to neural signals.

Is it always possible to manipulate a CDP with an SDP? One might think that if we take the survival-exemplary CDP and change a single SDP, we may – at least in theory – cause formal conflicts within the CDP. For example, suppose that we evaluate the survival-efficiency of an SDP for the physiological function A. Suppose that the feature value configuration that the SDP for A delivers varies with the feature value configuration for another physiological function B. Suppose also, conversely, that the feature value configuration that the SDP for B delivers varies with the feature value configuration for A. Let us suppose that the dependence between the SDP for A and the SDP for B is the following: for any value x taken by the SDP for B, the SDP for A also takes the value x, and for any value y taken by the SDP for A, the SDP for B takes the value $(y + 1)$. Here I assume that the values are measured using real numbers. In this example it may seem as we get a formal conflict, since there are no values for x and y such that $y = x$ and $x = y + 1$.

However, if we consider the example more closely, it should be clear that no formal conflict arises. An SDP describes how a trait token will *respond to* specific situations. The situation always precedes the response. This means that formal conflicts will never arise when manipulating a CDP with an SDP. In the above example, we do not get any formal conflict if we temporally separate the response from the situation. To see that no formal conflict arises, suppose that the situation is such that B is performed with the intensity 1 and A is also performed with the intensity 1. The trait token performing A will respond to this situation by continuing to perform A with intensity 1, whereas the trait token performing B will respond to this situation by performing B with intensity $1 + 1$ (i.e. 2). To this new situation, the trait token performing

A will respond by performing A with intensity 2, whereas the trait token performing B will respond by continuing to perform B with intensity 2, and so on. Hence, although performances of A and B are dependent on each other, no formal conflicts arise.

How does the manipulation of a survival-exemplary CDP with an SDP affect the CDP's further life expectancy? If the SDP that the survival-exemplary CDP is manipulated with is the same as the corresponding SDP in the original survival-exemplary CDP, then the further life expectancy remains the same. If the SDP differs, it might affect the further life expectancy. The different SDP may affect further life expectancy both by affecting the situation for other SDPs in the CDP, and by harming other SDPs (and thereby causing pathology of other trait tokens). Let us consider a concrete example. Say that we evaluate the survival-efficiency of the SDP for producing insulin of Ester's pancreas' system of beta cells. Suppose that Ester has type-1 diabetes mellitus. In cases of type-1 diabetes mellitus the beta cells in the pancreas get destructed. Normally, beta cells are responsible for producing insulin. A person with type-1 diabetes mellitus therefore becomes destitute of insulin. When we evaluate the survival-efficiency of Ester's SDP for producing insulin we consider this SDP in the context of an (otherwise) survival-exemplary CDP for Ester's reference class. That is, we consider the SDP as part of a survival-exemplary CDP manipulated with this SDP. The insulin deficit causes high blood sugar, which inter alia may lead to bodily responses such as fatigue, nausea, blurry vision, and weight loss. The high blood sugar may also affect further life expectancy by harming other traits, for example blood vessels and nerves. So, in the example, when we consider the survival-exemplary CDP of Ester's reference class manipulated with Ester's SDP for producing insulin, the further life expectancy will be affected both by changing the situation, which other SDPs react to, to include high blood sugar, and by the risks of impairments of the blood- and the nervous system's SDPs.

Let us now summarize how an SDP of a trait token is compared to the reference class standard. First, we separate the different SDPs for all of the trait token's type's physiological functions and make separate comparisons with the standard for all of them. The way we consider the SDP under evaluation in the comparison is as follows. We manipulate the survival-exemplary CDP of the reference class by substituting the SDP under evaluation for the corresponding SDP in the survival-exemplary CDP. We then consider the further life expectancy of the resulting CDP and compare that to the further life expectancy of the survival-exemplary CDP.

5.1.4 A Measure of Efficiency

Having specified the CDPs (complete disposition profiles) to compare with regard to their further life expectancy, we may now proceed to define 'survival-efficiency'. The question is now how to compare the two CDPs' further life expectancies.

A natural way to carry out the comparison is by division. We divide the further life expectancy of the survival-exemplary CDP manipulated with the SDP (specific disposition profile) of the trait token under evaluation by the further life expectancy of the exemplary CDP. As a result, we get a non-negative real number. This number is what I will term 'the survival-efficiency' of the evaluated trait token's SDP. Recall that $\boldsymbol{R}(O,t)$ is the set of all reference classes to which the organism O belongs at the time interval t, and that $\boldsymbol{S\text{-}Ex}(R)$ is the set of all survival-exemplary CDPs of the reference class R. Recall also that $fle(DP)$ is the further life expectancy of an organism with the CDP DP, and that $M(DP, dp)$ is the CDP DP manipulated with the SDP dp. Formally, we define 'the survival-efficiency of a trait token's SDP in the following way:

> Let O be an individual at the time interval t. Let dp be an SDP in the CDP representing O. Let $R \in \boldsymbol{R}(O,t)$ and let $DP_{S\text{-}Ex} \in \boldsymbol{S\text{-}Ex}(R)$. Provided that $fle(DP_{S\text{-}Ex}) > 0$, the **survival-efficiency** of dp relative to R and $DP_{S\text{-}Ex}$ is
>
> $$s\text{-}eff(dp, R, DP_{S\text{-}Ex}) = \frac{fle\big(M(DP_{S\text{-}Ex}, dp)\big)}{fle(DP_{S\text{-}Ex})}$$

If the survival-efficiency equals 1, it means that the SDP under evaluation meets the standard. A value < 1 means that the SDP does not meet the standard. And a value > 1 means that the SDP is above the standard. Note that the survival-efficiency of an SDP is undefined if the denominator equals 0.

We may note here that since there are possibly several reference classes to which an SDP's bearer belongs at the time of the evaluation, and also several survival-exemplary CDPs for each of those reference classes, there may be several different ways of calculating the survival-efficiency of an SDP. And there is a possibility that these yield different survival-efficiencies. When defining 'survival-health' and 'survival-pathology' in the second part of this chapter, all of these different ways will be taken into account. Until then, we may bracket this issue.

Let us, for illustrative purposes, make a concrete example of a survival-efficiency evaluation. Consider two human individuals belonging to the same

reference class, Frank and Georg. Frank and Georg are both 67 years old. We are interested in evaluating the survival-efficiency of Frank's and Georg's respective red bone marrow's SDPs for producing blood cells. Frank has some severe issues with his sight, but is otherwise like the survival-exemplary CDP of his reference class. Georg has acute myeloid leukemia, which means that his red bone marrow's capacity to produce certain blood cells (red blood cells, platelets, and some white blood cells) is severely reduced in comparison to the survival-exemplary CDP. Acute myeloid leukemia is a serious disease. Typical symptoms are tiredness, prolonged infections, fever, night sweats, and skeletal pain, and it takes longer time for the blood to clot in cases of injuries. If not treated, acute myeloid leukemia leads to death within some weeks or months (Regionala cancercentrum i samverkan 2014, 10).

Let us evaluate the survival-efficiency of Frank's and Georg's respective red bone marrow's SDPs for producing blood cells. For the purpose of this example, let us assume that Frank and Georg belong to one unique reference class at the time of the evaluation, and that there is one unique survival-exemplary CDP for this reference class. Given these assumptions, I will denote the survival-efficiency of Frank's, respectively Georg's, red bone marrow's SDP for producing blood cells '$s\text{-}eff(\text{Frank's red bone marrow's SDP})$' and '$s\text{-}eff(\text{Georg's red bone marrow's SDP})$'.

Suppose that the further life expectancy for the survival-exemplary CDP of Frank and Georg's reference class is 20 years. Then, since Frank's red bone marrow's SDP for producing blood cells is like the corresponding SDP in the survival-exemplary CDP, the further life expectancy of the survival-exemplary CDP manipulated with Frank's red bone marrow's SDP for producing blood cells is also 20 years. We may note here that, although Frank's further life expectancy may be lower than 20 years because of his sight issues, this does not affect the manipulated CDP, since this CDP is only manipulated with Frank's SDP for producing blood cells, not his SDP for seeing. Assuming these numbers we get:

$$s\text{-}eff(\text{Frank's red bone marrow's SDP}) = \frac{20}{20} = 1$$

Hence, Frank's red bone marrow's SDPs for producing blood cells meets the standard.

Let us then consider Georg. Since Georg has acute myeloid leukemia, his red bone marrow's SDP for producing blood cells is rather different from the corresponding SDP in the survival-exemplary CDP. Suppose that the further

life expectancy of the survival-exemplary CDP manipulated with Georg's red bone marrow's SDP for producing blood cells is 0,2 years, since this manipulated CDP represents an individual with acute leukemia. We then get:

$$s\text{-}eff(\text{Georg's red bone marrow's SDP}) = \frac{0{,}2}{20} = 0{,}01$$

Hence, Georg's red bone marrow's SDP for producing blood cells has a much lower survival-efficiency than Frank's. The survival-efficiency shows that the SDP is far from meeting the standard.

Let us now also define 'reproduction-efficiency' very similarly to how we define 'survival-efficiency' (recall that $\boldsymbol{R\text{-}Ex}(R)$ is the set of all reproduction-exemplary CDPs of the reference class R, and that $feno(DP)$ is the further expected number of offspring of an organism with the CDP DP):

> Let O be an individual at the time interval t. Let dp be an SDP in the CDP representing O. Let $R \in \boldsymbol{R}(O,t)$ and let $DP_{R\text{-}Ex} \in \boldsymbol{R\text{-}Ex}(R)$. Provided that $feno(DP_{R\text{-}Ex}) > 0$, the **reproduction-efficiency** of dp relative to R and $DP_{R\text{-}Ex}$ is
>
> $$r\text{-}eff(dp, R, DP_{R\text{-}Ex}) = \frac{feno\big(M(DP_{R\text{-}Ex}, dp)\big)}{feno(DP_{R\text{-}Ex})}$$

Since divisions where the denominator equals 0 are undefined, the above definitions of 'survival-efficiency' and 'reproduction-efficiency' will not give us any efficiency value if the denominator equals 0. Is it possible for the denominator in the above definitions to equal 0? Yes, at least with regard to reproduction-efficiency, that should be possible. Here, it is not just in principle possible, but the case for some reference classes. Consider for example mules, or elderly human females. Since it is not readily attainable for any mule to be able to get offspring, the further expected number of offspring of a survival-exemplary CDP for a reference class of mules will be 0. Similarly for elderly human females. With regard to survival-efficiency, it is questionable whether it is in principle possible for the denominator to equal 0. If the further life expectancy is 0, this should mean that the organism is already dead, since if it is expected to die in the next time interval, there is still one time interval left.

Although there may be reference classes where the reproduction-exemplary CDP's further expected number of offspring is 0, it is perhaps not so much of a problem that the definition of 'reproduction-efficiency' leaves such cases

undefined. If the reproduction-exemplary CDP's further expected number of offspring is 0, then it should not be relevant to consider the reproduction-health status of any trait token among the members of the reference class.

However, if we would like to define the reproduction-efficiency, or even the survival-efficiency, also in cases where the denominator equals 0, we could add the following conditions to the definitions:

> If both the denominator and the numerator equal 0, then let the efficiency be 1.
>
> If the denominator equals 0 and the numerator takes a value higher than 0, then let the efficiency be > 1.

This means that in all cases where the denominator equals 0, the trait token under evaluation will be healthy (given that there are no efficiency reductions for some other SDP of the trait token). This, I think is in line with the reasoning underlying the above presented efficiency models – if the standard to which we compare the SDP of a trait token is not better with regard to survival (or, respectively, with regard to reproduction) than the SDP under evaluation, then there is no efficiency reduction. If both the denominator and the numerator equal 0, then I think it is reasonable to say that the efficiency is 1, since the two further life expectancies that we compare are the same. However, if it is the case that the numerator takes a value higher than 0 when the denominator equals 0, then it is a bit strange to say that the efficiency is 1, since the SDP of the trait token under evaluation is better off with regard to survival (or, respectively, reproduction) than the standard. Hence, in such cases the efficiency should be higher than 1. But since we cannot get a more precise result by division in this case, we will have to be content with the result that the efficiency is > 1. Unfortunately, this means that we cannot compare the efficiencies on a quote scale in such cases.

5.2 Health and Pathology

Having defined 'survival-efficiency' and 'reproduction-efficiency', let us consider the connection between these efficiency measures and survival-health and -pathology and reproduction-health and -pathology. Roughly, the idea is that a trait token is survival-pathological if and only if at least one of its SDPs (specific disposition profiles) for the physiological functions for its type has a survival-efficiency that is too low. And, if and only if all of the trait

token's SDPs for the physiological functions for its type have survival-efficiencies that are sufficiently high, the trait token is survival-healthy. Mutatis mutandis for reproduction-health and -pathology.

In this section, I will define 'survival-health' and 'survival-pathology', and respectively 'reproduction-health' and 'reproduction-pathology', of a trait token in two steps. First, I present premature definitions, which deal with the issue of how to draw the line between too low efficiencies and high enough efficiencies (Section 5.2.1). Then, I add some complexity to these definitions (Section 5.2.2). The added complexity is needed to take into consideration that efficiencies are relative to a choice of reference class and a choice of survival- or reproduction-exemplary CDP (complete disposition profile).

5.2.1 Premature Definitions

As already indicated, the survival-health of a trait token is a relation between the SDPs (specific disposition profiles) of the trait token and its reference class. The thought is that this relation is the same irrespectively of which trait type is being considered. Therefore, what counts as a too low survival-efficiency, and, respectively, a high enough survival-efficiency, should be the same irrespectively of the type of trait token that is evaluated, and irrespective of the reference class of the trait token's bearer. For this reason, the line distinguishing survival-health from survival-pathology will be represented by a constant. Survival-efficiencies taking the value of this constant, and survival-efficiencies taking values below this constant, will indicate survival-pathology, whereas survival-efficiencies taking values higher than this constant will indicate survival-health.

In the following premature definitions of 'survival-health' and 'survival-pathology', let us assume that it is always the case that the organism with the trait token under evaluation belongs to one unique reference class at the time of the evaluation and that there is one unique survival-exemplary CDP (complete disposition profile) for this reference class. Let k_s be a constant real number that represents the described distinguishing line. We then, prematurely, define 'survival-health' and 'survival-pathology' as follows:

> A trait token a is **survival-healthy** if and only if each of its SDPs for the physiological functions for a's trait type has a survival-efficiency above k_s.

A trait token a is **survival-pathological** if and only if at least one of its SDPs for a physiological function for a's trait type has a survival-efficiency equal to or below k_s.

We may similarly define 'reproduction-health' and 'reproduction-pathology'. Let k_r be a constant that represents the line distinguishing too low reproduction-efficiencies from high enough reproduction-efficiencies. We prematurely define 'reproduction-health' and 'reproduction-pathology' as follows:

A trait token a is **reproduction-healthy** if and only if each of its SDPs for the physiological functions for a's trait type has a reproduction-efficiency above k_r.

A trait token a is **reproduction-pathological** if and only if at least one of its SDPs for a physiological function for a's trait type has a reproduction-efficiency equal to or below k_r.

For illustration of these definitions, remember the above example about Frank and Georg. We evaluated the survival-efficiency of their respective red bone marrow's SDP for producing blood cells. We saw that Frank, whose red bone marrow's SDP for producing blood cells was like that of the survival-exemplary CDP, had the survival-efficiency 1. And we saw that Georg's red bone marrow's SDP for producing blood cells had the survival-efficiency 0,01, because of Georg's acute myeloid leukemia. If we now proceed to evaluate the health status of Frank's and Georg's respective red bone marrows, what we need to consider is whether these respective values are higher, equal to, or lower than k_s. If the value is higher, and if there is no other physiological function for red bone marrows, for which the individual's red bone marrow have SDPs with survival-efficiencies equal to, or below, k_s, then the red bone marrow is survival-healthy. If the value is not higher than k_s, then the red bone marrow is survival-pathological.

Akin to how Boorse defines the biostatistical theory's distinguishing line, I take the values of k_s and k_r to be conventional, but within certain limits. The values of k_s and k_r should be smaller than 1. Otherwise we will have pathology without any reduction in efficiency. Hence, in the above example, the survival-efficiency of Frank's red bone marrow's SDP for producing blood cells indicates that the red bone marrow is survival-healthy. Definitely for k_s, and reasonably also for k_r, the value should also be greater than 0. Otherwise,

for a trait token to be survival-pathological, it would be required that the further life expectancy of the survival-exemplary CDP manipulated with some of the trait token's SDPs is as low as 0. And, for a trait token to be reproduction-pathological, it would be required that the reproduction-exemplary CDP manipulated with some of the trait token's SDPs has no further expected number of offspring. Further, the values of k_s and k_r are both reasonably closer to 1 than 0. If k_s were closer to 0, then a state would have to reduce further life expectancy to a very large extent for it to count as survival-pathological. If we consider the above example, the survival-efficiency of Georg's red bone marrow's SDP for producing blood cells is 0,01. Reasonably, the value of k_s is at least closer to 1 than that, and hence Georg's red bone marrow is survival-pathological.

That k_s and k_r are constants means that k_s and k_r are to be used irrespectively of which SDP is evaluated, although SDPs for different physiological functions may be more or less important for survival; the same holds for reproduction. So, for example, regardless of whether we evaluate the survival-efficiency of a token heart's SDP for pumping blood, or a token melanocyte's SDP for producing melanin (pigments protecting the human body from UV-radiation), the survival-efficiency distinguishing survival-health from survival-pathology is k_s. Since hearts are more important for survival than single melanocytes, this means that it will be more likely for hearts than for single melanocytes to be pathological. If a heart token is not perfectly disposed to pump blood, this may have significant negative effects on further life expectancy, which means that the efficiency will be significantly reduced. In contrast, even if a single melanocyte is functioning quite far from perfectly, this may barely at all affect further life expectancy, which means that the efficiency will be very close to 1. In Chapter 7 (Section 7.1), I will discuss implications related to this feature of the disposition profile efficiency theory.

5.2.2 Final Definitions

Let us add the remaining complexity to the above definitions. Since the survival-efficiency of an SDP (specific disposition profile) of a trait token is relative to a choice of reference class and survival-exemplary CDP (complete disposition profile), we cannot just judge the trait token survival-healthy based on one choice of reference class and survival-exemplary CDP.

What should we do, then, if the organism with the trait token that we evaluate belongs to several reference classes at the time of the evaluation, and if there are several survival-exemplary CDPs for these reference classes? The

idea is the following: if for all possible choices of reference class and survival-exemplary CDP, we get the result that the survival-efficiency of a particular SDP of a trait token is below or equals k_s, then the trait token is survival-pathological. If for all possible choices of reference class and survival-exemplary CDP, we get the result that the survival-efficiency for each of the trait tokens SDPs is above k_s then the trait token is survival-healthy. If for some choices of reference class and survival-exemplary CDP, we get the result that the survival-efficiency of a particular SDP of a trait token is below or equals k_s, and for some choices get the result that the survival-efficiency of this SDP is above k_s, then it is indeterminate what the survival-health status of the trait token is.[27]

Remember that $\boldsymbol{R}(O,t)$ is the set of all reference classes to which the organism O belongs at the time interval t, and that $\boldsymbol{S\text{-}Ex}(R)$ is the set of all survival-exemplary CDPs of the reference class R. Formally, we define 'survival-health' and 'survival-pathology' as follows:

> A trait token a of an organism O is **survival-healthy** at a time interval t if and only if each of its SDPs dp for the physiological functions for a's trait type is such that, for each $R \in \boldsymbol{R}(O,t)$ and each $DP_{S\text{-}Ex} \in \boldsymbol{S\text{-}Ex}(R)$, $s\text{-}eff(dp, R, DP_{S\text{-}Ex}) > k_s$.

> A trait token a of an organism O is **survival-pathological** at a time interval t if and only if at least one of its SDPs dp for a physiological function for a's trait type is such that, for each $R \in \boldsymbol{R}(O,t)$ and each $DP_{S\text{-}Ex} \in \boldsymbol{S\text{-}Ex}(R)$, $s\text{-}eff(dp, R, DP_{S\text{-}Ex}) \leq k_s$.

For illustration, let us return to the example about Frank and Georg. If Frank belongs to more than one reference class at the time of the evaluation, and if there for some of these reference classes are more than one survival-exemplary CDP, then all of this must be taken into consideration in the survival-health evaluation. Frank's red bone marrow will be survival-healthy only if it, for any possible choice of reference class that Frank belongs to at the time of the evaluation, and any possible choice of survival-exemplary CDP for these

[27] This is reminiscent of supervaluationism in logic. According to supervaluationism, a proposition is true if and only if it comes out true under all precisifications, and a proposition is false if and only if it comes out false under all precisifications, and a proposition lacks truth-value if and only if there is some precisification under which the proposition comes out true and some precisification under which the proposition comes out false. (See for example Sorensen (2018).)

reference classes, holds that the survival-efficiency is above k_s. And Frank's red bone marrow will be survival-pathological only if for each of these choices it holds that the survival-efficiency equals or is below k_s.

Similarly to the above definitions, we may define 'reproduction-health' and 'reproduction-pathology' (remember here that $\boldsymbol{R\text{-}Ex}(R)$ is the set of all survival-exemplary CDPs of the reference class R):

> A trait token a of an organism O is **reproduction-healthy** at a time interval t if and only if each of its SDPs dp for the physiological functions for a's trait type is such that, for each $R \in \boldsymbol{R}(O,t)$ and each $DP_{R\text{-}Ex} \in \boldsymbol{R\text{-}Ex}(R)$, $r\text{-}eff(dp, R, DP_{R\text{-}Ex}) > k_r$.

> A trait token a of an organism O is **reproduction-pathological** at a time interval t if and only if at least one of its SDPs dp for a physiological function for a's trait type is such that, for each $R \in \boldsymbol{R}(O,t)$ and each $DP_{R\text{-}Ex} \in \boldsymbol{R\text{-}Ex}(R)$, $r\text{-}eff(dp, R, DP_{R\text{-}Ex}) \leq k_r$.

Is the efficiency measures' relativity to a choice of reference class and survival- or reproduction-exemplary CDP a problem? If it is the case that there are several possible choices of reference class and survival- or reproduction-exemplary CDP for the efficiency evaluation of most SDPs, and these different choices also yield efficiencies differing in whether they are lower than or equals k_s/k_r or whether they are higher than k_s/k_r, then the above account will judge most trait tokens' health statuses indeterminate. That would obviously be problematic: first, since the account would conflict with many classifications in physiology and pathology of states as healthy or pathological, and, second, since an account that says of most states that it is indeterminate whether they are healthy or pathological seems rather unfruitful.

However, the theoretical possibility of there being many possible choices of reference classes and survival- or reproduction-exemplary CDPs, yielding very different efficiencies, is not necessarily a problem. I will argue that, for most SDPs, we should not expect there to be several choices that also yield efficiencies differing in whether they are below or equal to k_s/k_r, or are above k_s/k_r.

Considering reference classes it should, as pointed out in Chapter 4 (Section 4.2), because of the biological world being as it is, be rare that an organism belongs to more than one reference class at the time of the evaluation. But, more importantly, given how the biological world actually is, if an organism

belongs to more than one reference class at the time of the evaluation, these reference classes are probably rather similar.

Similarly, considering survival-exemplary CDPs, we should, as I also argued in Chapter 4 (Section 4.1), expect that if a reference class has several survival-exemplary CDPs, these are rather similar. Mutatis mutandis for reproduction-exemplary CDPs.

In light of this, we may expect it to be unusual that there are several very different reference classes to which the organism with the trait token under evaluation belongs at the time of the evaluation and that there are many very different survival- or reproduction-exemplary CDPs for the different reference classes. This means that even if there are several reference classes to which the organism with the trait token under evaluation belongs at the time of the evaluation and there are several survival- or reproduction-exemplary CDPs for these reference classes, this need not be a serious problem. Since the different alternatives should be rather similar, they should not give rise to large efficiency variations. This means that most trait tokens have a determinate survival- or reproduction-health status. A case where it is indeterminate whether a trait token is survival- or reproduction-healthy or -pathological should typically be a case where the different efficiencies are all close to k_s or k_r but some are just above and some equal to or below k_s or k_r.

If we again consider the example about Frank and Georg, it seems that the result will be that Frank's red bone marrow is survival-healthy and Georg's red bone marrow is survival-pathological. Even if Frank belongs to several reference classes at the time of the evaluation, and there are several survival-exemplary CDPs for these reference classes, each choice of reference class and survival-exemplary CDP will presumably result in very similar survival-efficiencies around 1, i.e. well above k_s. And even if there are several choices of reference class and survival-exemplary CDP for the evaluation of Georg's red bone marrow's SDP for producing blood cells, each of these choices will presumably result in very similar survival-efficiencies around 0,01, i.e. well below k_s.

6 The Problem of Common Diseases and Kingma's Dilemma

In Chapter 2 (Sections 2.3.3-2.3.4), I described two objections raised against the BST which have gained a lot of attention. One is the problem of common diseases, and the other is Kingma's dilemma. In this chapter I will discuss earlier suggested solutions to these two problems and explain how the disposition profile efficiency theory solves them. I will first consider the problem of common diseases (Section 6.1) and then Kingma's dilemma (Section 6.2).

6.1 Solving the Problem of Common Diseases

The problem of common diseases was described in Chapter 2 (Section 2.3.4) as a serious problem for the biostatistical theory. Boorse distinguishes health from pathology by ordering the actual efficiencies in the reference class, and drawing the distinguishing line a certain distance below the average efficiency within this distribution. This line-drawing makes it problematic for Boorse to classify states of common diseases as pathological. Rather, specific types of pathological conditions are, on the BST, uncommon by definition. Examples of common diseases, which I have mentioned, are Alzheimer's disease (among elderly humans), canin hip dysplasia (among some dog breeds), benign prostatic hypertrophy (among elderly men), and dental caries (among humans).

Beneath, I will show that the disposition profile efficiency theory solves the problem of common diseases. However, the DPE-theory is not the first theory to suggest a solution to the problem of common diseases. Before explaining how the DPE-theory solves the problem (Section 6.1.5), I will discuss earlier suggestions of how to solve the problem by Kraemer (2013) (Section 6.1.1), Schwartz (2007) (Section 6.1.2), Hausman (2012) (Section 6.1.3), and Garson and Piccinini (2014) (Section 6.1.4).

6.1.1 The Statistical Fitness Theory

Kraemer (2013) suggests his "statistical fitness theory" (abbreviated 'SFT') as a solution to the problem of common diseases. We should note, though, that the SFT does not distinguish health from pathology, but "normal function" from "malfunction" or "dysfunction".

There are three aspects of the SFT that are supposed to do the job of solving the problem of common diseases. These aspects regard, first, the individuation of reference classes; second, the strength required for physiological function; and, third, the line-drawing between normal function and malfunction.

Kraemer argues that reference classes should not be individuated, as Boorse suggests, by sex and age, but only by species and sometimes also geographical isolation (2013, 429-430). Normal typical differences between sexes and ages, Kraemer thinks, can instead be taken care of by a "certain circumstances clause". Kraemer writes:

> We do better to leave the reference class as a whole species and index functions to certain circumstances that take into account such factors as development and environment. That is SFT's strategy. (2013, 430)

Regarding the strength required for physiological functions (or, in Kraemer's terminology, simply 'functions'), Kraemer is very precise. He writes:

> SFT makes this notion [of having a function] more precise than previous statistical theories of functions do. It holds that when *x*'s typically do *y* in a reference class under certain circumstances, that entails that roughly 30 % or more of the *x*'s in the reference class do *y* at some point, under those circumstances. If fewer than 30 % of the *x*'s do *y* in a reference class under certain circumstances, then *x*'s do *not* have the function to *y* in that reference class under those circumstances. (Kraemer 2013, 428, italics are original)

Kraemer then accounts for malfunction/dysfunction as follows:

> Where *x* refers to a token of a type of trait and *r* refers to a reference class,
>
> An *x* in *r* under certain circumstances malfunctions if *x* has the function to *y* in *r* under those circumstances and *x*'s *y*-ing in those circumstances falls below the normal range of the performance *y* by *x*'s in *r* under those circumstances. (2013, 430)

This looks very similar to Boorse's line-drawing between health and pathology. However, there are two important differences. First, by "the normal range of the performance" Kraemer does not, as Boorse, refer to an ordering "according to efficiency at contributing to survival and reproduction" (Kraemer 2013, 431). Rather, he refers to an ordering "by the *actual effects* of the token traits" (Kraemer 2013, 431, italics are original). So, for example, if we evaluate a token thyroid's secretion of the hormones T3 and T4, we should not consider to what extent its secretion of these hormones contributes to survival or reproduction. Rather, we should merely consider the amount of T3 and T4 that it secretes. Second, Kraemer is more precise than Boorse about when a performance falls below the normal range. He writes:

> According to SFT, the *normal* range of effects of a trait is clustered around a weighted mean. The normal range is roughly up to two standard deviations above and below the weighted mean. Malfunctioning trait tokens fall two standard deviations below the weighted mean and further below that. Any token of a trait, including tokens with zero *y* effect, that falls below the normal range for that trait is designated as malfunctioning by SFT. (Kraemer 2013, 431, italics are original)

It should be clarified that the SFT's "normal range of effects" does not include trait tokens that have no activity at all. Kraemer writes: "SFT […] determines normal functioning by using an ordering of those traits that actually have certain effects, which typically contribute to fitness" (2013, 434). Hence, the line distinguishing malfunctioning trait tokens from normally functioning trait tokens only takes into consideration those trait tokens that perform physiological functions to some degree. Nevertheless, the line applies also to those trait tokens that have no activity at all.

Let us now consider whether the SFT solves the problem of common diseases. The way the SFT is supposed to solve this problem is, first, by viewing whole species as reference classes. The thought is that, because of this choice, many conditions described as common diseases because they are common within a certain age group (e.g. Alzheimer's dementia) cannot be said to be common in the reference class. Second, the SFT allows that roughly 70 percent of the tokens of a trait type in a reference class are malfunctioning by two means: (i) it defines the line between function and malfunction with reference to the distribution of the effects only of those trait tokens that have some activity, and (ii) it only requires that 30 percent of the tokens in the reference

class perform the function under certain circumstances in order for the function to count as a physiological function for the trait type. I will explain in more detail how this is supposed to work out below.

I will argue that the SFT is not able to account for most cases of common diseases. What we should first note is that the choice to view whole species as reference classes does not contribute to solving the problem of common diseases. Since functions are instead indexed to circumstances including both environment and development, diseases that are common, for instance, in certain age groups will be common in the whole reference class – given a certain stage of development. For example, it is typical for humans to have a reduced memory capacity in the developmental stage that occurs around the age of 90 years.

Second, the SFT allows for up to just over 70 percent of the tokens of a trait type in a reference class being dysfunctional. The explanation for this is as follows. On the SFT, something counts as a physiological function only if at least around 30 percent of the reference class members are disposed to perform the function in a certain situation. The line distinguishing dysfunctional trait tokens from normally functioning trait tokens is based only on those trait tokens that are disposed to have some activity of the function. In cases where only 30 percent of the trait tokens are disposed to have some activity, the line distinguishing dysfunctional trait tokens from normally functioning trait tokens will be based only on those 30 percent. The line will be drawn roughly two standard deviations below the average quantitative activity for these 30 percent of trait tokens, leaving about the 2,5 percent of these 30 percent with the lowest activity dysfunctional. Since the same line determines the functional status of the trait tokens that are not disposed to have any activity for the concerned physiological function, all of these trait tokens are dysfunctional as well – simply because they have a lesser activity than the activity at two standard deviations below the average activity for the tokens that have some activity. In this case this amounts to 70 percent of the reference class members. In sum, then, just over 70 percent of the trait tokens are dysfunctional in a case where only 30 percent of the reference class members are disposed to have some activity of the physiological function. So, if common diseases are such that they tend to make concerned individuals unable to perform the physiological function in question by any activity whatsoever, then the SFT get sound results in cases of common diseases. However, common diseases are typically not like that. If we, for example, consider the capacity to memorize, this capacity becomes significantly reduced by Alzheimer's disease, however, typically not completely lost. The same goes for canine hip

dysplasia in dogs. This condition reduces the functionality of the hip, however, it does not completely take its functionality away. And if we consider benign prostatic hypertrophy, this condition typically makes it difficult to urinate, but it does typically not make it impossible to urinate. As long as the function, for most trait tokens affected by the condition, is not lost but merely severely reduced, the SFT does not classify more than around 2,5 percent of the trait tokens in the reference class dysfunctional. In these cases, the distribution of quantitative performances will include most trait tokens in the reference class, and the line at two standard deviation below the average quantitative performance will leave about 2,5 percent of more or less the whole reference class dysfunctional. We may hence conclude that the SFT does not solve the problem of common diseases.

6.1.2 The Frequency and Negative Consequences Approach

Schwartz (2007) adds an aspect of negative consequences to the biostatistical theory in order to solve the problem of common diseases. He calls the resulting theory "the frequency and negative consequences approach". As already described in Chapter 2 (Section 2.3.2) and Chapter 3 (Section 3.1.1), Schwartz takes the relevant negative consequences to be consequences that are negative for the carrying out of activities that are standard in the species and have been for a long period of time. When discussing what the relevant goals for physiological functions, health, and pathology should be in Chapter 3 (Section 3.1.1), I argued against adding the aspect of negative consequences that Schwartz proposes. Let us here bracket that discussion and let us only focus on whether Schwartz's added aspect of negative consequences solves the problem of common diseases.

The way Schwartz thinks that negative consequences should be taken into account in health status evaluations is as an additional *Z*-axis, besides the "the *X*-axis showing the level of functioning" and "the *Y*-axis showing prevalence" (2007, 377). This *Z*-axis he thinks should affect the line-drawing in the following way:

> More severe negative consequences would support moving the line between normal and abnormal (on the *X*-axis) toward the *right*, to count more levels of functioning as dysfunctional. A lack of negative consequences would support moving the line to the *left*, to count only lower levels of functioning (or even no levels of functioning at all) as dysfunctional. (Schwartz 2007, 377, italics are original)

Schwartz exemplifies how the addition of negative consequences solves the problem of common diseases as follows:

> Consider the problem of common disease, such as BPH [benign prostatic hypertrophy] in 70-year-old men or Alzheimer's in 85-year-old men or women. In these cases, the negative consequences would support moving the line between low-normal and dysfunctional on the *X*-axis to the right. Similarly, the serious effect of canine hip-dysplasia on ambulation would support considering the lowest 30% of the breed as having dysfunctional hips. (2007, 377)

Although Schwartz's suggestions may at first sight seem to solve the problem of common diseases, I will argue that it does not. A problem with Schwartz's suggestion is that he does not discuss any comparison level for negative consequences. This is problematic, because in order to know whether a state should count as having negative consequences, say, for the ability to see, we need to know what the comparison level for seeing is: in comparison to what ability to see does the state have negative consequences? For human eyes, this level cannot reasonably include seeing as sharp as an eagle, since this would mean that the states of all human eyes have negative consequences. What it seems Schwartz (implicitly) has in mind is something along the following lines: the comparison level is set by what is typical for the species, or perhaps what is typical for the best functioning reference classes of the species. Adopting this comparison level, however, results in very strange disease ascriptions. For example, Alzheimer's disease makes one worse than the average human at speaking or running. However, having Alzheimer's disease is not the only possible reason for being worse at speaking or running. Other possible reasons are babyhood and childhood. But we do not want to say generally that the typical level of speaking or running abilities in babies and children have negative consequences that should affect our health or pathology judgements about them. Hence, for Schwartz's suggestion to make sense, he cannot assume that the comparison level for negative consequences is determined by what is typical for the whole species, or the best functioning reference classes in the species. Rather, the comparison level must be relativized to reference classes, which are more fine-grained than encompassing whole species. However – and here we get a dilemma – if we make the comparison level relative to more fine-grained reference classes we will be back on square one with the problem of common diseases. If it is typical in a reference class to have a state with certain consequences, and the comparison level is relative to the reference class, these consequences cannot count as negative. Hence, diseases that

are common in a reference class will not count as having negative consequences. For example, if Alzheimer's disease is very common in a reference class, this disease will influence the standard for performing activities including speaking and running, which means that Alzheimer's disease will not count as a state with negative consequences.

An alternative way to determine the levels for negative consequences would be with reference to what is valued for humans. If we understand Schwartz as intending this understanding of negative consequences, he does not end up in the dilemma presented above. However, if one understands negative consequences as determined by human values, the theory will obviously involve values, which I put as a desideratum in Chapter 1 that the sort of theory that this thesis concerns must not do. Schwartz does not explicitly express that the theory he aims at should involve no values. However, since he presents his theory as a further development of Boorse's theory, it is natural to conceive him as intending a value-free theory.

I think that what the above discussion indicates is that what we need in order to solve the problem of common diseases is not, in the first instance, a change in what sorts of facts to take into consideration in health status evaluations. Rather, what we need is an alternative idea to Boorse's of how to determine the comparison level, or standard, for health, and how to draw the line separating health from pathology. This is, as we will see below, precisely what Hausman proposes.

6.1.3 The Functional Efficiency Theory

I have described Hausman's "functional efficiency theory" both in Chapter 4 (Section 4.1.1) and Chapter 5 (Section 5.1.2). This theory, Hausman claims, solves the problem of common diseases. As I understand Hausman's theory, it solves the problem by two means. The first concerns how the standard for health is defined, and the second concerns the theory's view on the distinction between too low efficiencies and high enough efficiencies.

We saw in Chapter 2 (Section 2.2.4) that Boorse accounts for the standard for health in terms of averages in the reference class. This makes it impossible for the biostatistical theory to classify any condition that affects half, or more, of the reference class as pathological. In contrast, Hausman accounts for the standard for health in terms of what is readily attainable for the reference class members. This makes it possible to classify conditions that affect more than half of the reference class members as pathological in cases where there is a

discrepancy between the actual capacities of the reference class members and the best capacities readily attainable for them.

We also saw in Chapter 2 (Section 2.2.4) that, although Boorse is not clear on precisely how the line distinguishing high enough efficiencies from too low efficiencies is to be drawn, at least he seems to define the distinguishing line with reference to the distribution of efficiencies in the reference class. He thinks that the line should be drawn below, or far below, the average efficiency level. This line-drawing further limits how common a disease may be. According to this line-drawing, a specific type of pathological state cannot be more common than affecting at maximum a little less than half, or significantly less than half, of the reference class members. Hausman argues that the prevalence of different efficiency levels should not determine the line distinguishing health from pathology. He writes:

> […] the functional efficiency theory, in contrast to the BST, denies that prevalence defines whether a level of functional efficiency is healthy or pathological. Although what is prevalent is typically healthy, the problems of common diseases and healthy populations show that this coincidence is not perfect. (Hausman 2012, 536)

When describing the distinguishing line of the functional efficiency theory, Hausman then writes:

> First, if a level of functional efficiency is "maximal"–that is, at or above one of the highest levels readily attainable in some benchmark environment–then it is healthy. Second, if a level of functional efficiency is significantly worse than a maximal level, then it is pathological. (2012, 537)

However, as noted in Chapter 5 (Section 5.1.2), Hausman's definition of 'efficiency' merely accounts for efficiency comparatively. This allows us to order the efficiencies of token capacities; however, I argued, it does not allow us to draw a general line representing what is significantly worse than the standard that is not determined by the distribution of efficiencies.

We may conclude that Hausman's ideas for solving the problem of common diseases are promising. However, because of struggles with the notion of efficiency, the functional efficiency theory does not reach all the way to solve the problem.

6.1.4 A Biostatistical Account of Functions

Garson and Piccinini suggest a biostatistical account of functions, which they present as "a novel and improved version of Boorse's biostatistical theory of functions" (2014, 1). Similarly to Kraemer (2013), Garson and Piccinini's account does not distinguish health from pathology, but normal function from dysfunction.

Garson and Piccinini first define 'physiological functions' (or, in their terminology, simply 'functions') (2014, 6):

> Formally, we can define the function of a trait in terms of a conditional probability. If *X* is a trait, *Y* is an activity of that trait, and *RC* is the reference class, then:
>
> A function of *X* in *RC* is $Y =_{\text{def}}$
>
> (i) *P*(*X* contributes to survival or inclusive fitness in *RC*) is non-negligible.
> (ii) *P*(*X* contributes to survival or inclusive fitness in *RC* by doing $Y \mid X$ contributes to survival or inclusive fitness in *RC*) is non-negligible.

Garson and Piccinini then define 'situations that are appropriate for the performance of a function' (2014, 8):

> To identify situations as appropriate or inappropriate for the exercise of a function, we can use conditional probability again. Let *RC* be a reference class, *X* a trait that is distributed over that reference class, and *Y* a function of that trait. Let *S* be a situation-type (i.e. a set of instances of situations), such as asphyxiation due to an obstructed throat, or a disjunction of such situation-types, such as asphyxiation due to an obstructed throat or ingestion of a toxic substance. Relative to *RC*, the set of situations, *S*, under which *X*'s exercise of *Y* is appropriate can be defined by the following two (independently necessary and jointly sufficient) conditions:
>
> (i) Inclusivity: *P*(*X* is in *S* | *X*'s doing *Y* contributes to survival or inclusive fitness) ≈1.
> (ii) Specificity: there is no *S′* which is a proper subset of *S* such that (i) is true of *S′*.

In order to account for the dysfunction, not only of trait tokens that have absolutely no functional activity in some appropriate situation, but also of trait tokens that have some – although not an appropriate – functional activity, Garson and Piccinini define 'appropriate rates of functioning' (2014, 10):

> Let *RC* be a reference class, *X* a trait that is distributed over that reference class, *Y* a function of that trait, and *S* the set of situations appropriate for the performance of *Y* by *X*:
>
> Trait *X* performs function *Y* at a rate of functioning that is appropriate in a situation $s =_{def}$
>
> (i) If an organism in *RC* possessing *X* is in *s* and $s \notin S$, then *X* performs function *Y* at a rate of zero (or close to zero).
> (ii) If an organism in *RC* possessing *X* is in *s* and $s \in S$, then *X*'s rate of functioning provides an adequate contribution to survival or inclusive fitness in *s*, relative to other rates that are physiologically possible for *X* in *s*.

Garson and Piccinini then define 'dysfunction' as follows (2014, 13):

> *X* is dysfunctional with respect to $Y =_{def}$.
>
> *X* cannot perform *Y* in at least one situation, *s*, with $s \in S$, at the rate that is appropriate in *s*.

Although Garson and Piccinini's account looks rather different from Hausman's, the idea for solving the problem of common diseases is similar. We saw that Hausman comes close to solving the problem of common diseases by two means. First, by accounting for the standard for health in terms of what is readily attainable for members of the reference class, and, second, by denying that the line between health and pathology is dependent on the distribution of efficiencies in the reference class. In Garson and Piccinini's account, there are similar ideas.

First, in the definition of 'appropriate rates of functioning' Garson and Piccinini express a standard for health in terms of what is physiologically possible for the tokens of a trait type in certain situations. To illustrate the idea of physiological possibility, Garson and Piccinini use an example about frogs:

> For example, one explanation of the rapid decline of the Cascades frog, *Rana cascadae*, in Oregon appeals to the recent increase in ultraviolet-B radiation, which appears to affect immune response and hence render the larvae more susceptible to a regional fungus (Sakar [1996]). Regardless of whether this immune response becomes pandemic, the frog's lowered immune response does not provide an adequate contribution to survival or inclusive fitness when it

> leads to premature death of larvae. This is so because the lowered immune response leads to a substantial fitness decrease relative to what we know to be physiologically possible for the frog. (2014, 11-12)

Although Garson and Piccinini do not explain in detail how to cash out "physiologically possible", their idea is reminiscent of Hausman's idea of what is readily attainable for a nonnegligible number of members in the reference class in common environments of the reference class. In contrast to Boorse, Garson and Piccinini, just like Hausman, do not define the standard in terms of actual averages in the reference class but rather in terms of what could be expected of members of the reference class.

Second, Garson and Piccinini count every trait token that does not perform its physiological functions at the appropriate rate in each situation dysfunctional. Hence, when drawing the line between normal function and dysfunction they, like Hausman, do not refer to the distribution of different functionings in the reference class.

We may conclude that Garson and Piccinini's biostatistical account of functions is promising for solving the problem of common diseases. However, their account only defines 'dysfunction', and not 'pathology'. According to Garson and Piccinini's suggestion, every trait token that is not disposed to function at the level of the comparison standard is dysfunctional. This makes their notion of dysfunction different from how most conceive of pathology. Usually, pathology is taken to require not just any reduction, but a sufficient reduction, from the comparison standard. Garson and Piccinini's notion of dysfunction does thereby not enable comparison of efficiencies between non-perfectly functioning trait tokens. So, in contrast to the disposition profile efficiency theory which allows for comparisons of how well different trait tokens are disposed to perform their physiological functions on a ratio scale, and also in contrast to Hausman's functional efficiency theory which at least allows for comparisons of how well different trait tokens are disposed to perform their physiological functions on an ordinal scale, Garson and Piccinini's account count every trait that is not perfectly functioning dysfunctional.

6.1.5 The Disposition Profile Efficiency Theory

To a large extent, the disposition profile efficiency theory follows Hausman's ideas for solving the problem of common diseases; that is (1) by defining a standard for health distinct from the actual average in the reference class, and (2) by defining the line distinguishing health from pathology without reference

to the distribution of efficiencies in the reference class.

By accounting for the standard for health in terms of what is readily attainable for the members of a reference class, it becomes possible for the DPE-theory to classify states that affects more than half of the reference class members as pathological when the states of the reference class members are not as beneficial for high further life expectancy, or high further expected number of offspring, as what is readily attainable for them. In relation to Hausman's functional efficiency theory and Garson and Piccinini's biostatistical account of function, a benefit of the DPE-theory's account of the standard is that it is more precise. An important notion in Garson and Piccinini's account of the standard is that of being physiologically possible. However, Garson and Piccinini do not fully articulate what it means that something is physiologically possible for the tokens of a trait type. An important notion in Hausman's account, which is not fully articulated, is that of being readily attainable. The DPE-theory's standard is reminiscent of Hausman's account of the standard; however, the notion of being readily attainable is defined more clearly.

By defining the distinguishing line merely with reference to a relation to the standard, the distinguishing line does not impose any limitation on how common a certain pathological state may be. With regard to this part of the solution to the problem of common diseases, we find important differences between the DPE-theory, on the one hand, and Hausman's and Garson and Piccinini's theories, on the other. We saw that Garson and Piccinini's theory can only distinguish between a trait token working exactly as well as the standard and a trait token not working exactly as well. It is not able to account for more nuances in how well a trait token is functioning. Hausman's theory allows for some more nuances. It can tell us whether a trait token works better or worse than, or equally well as, another trait token of the same type (e.g. the standard for that trait token). However, the comparisons that can be made are rather limited. Although the theory can order trait tokens with regard to how well they function, it cannot tell how much better or worse a certain trait token works than another trait token. Hence, Hausman's theory does not allow us to say that pathology is about functioning so and so much worse than the standard, at least if "so and so much" is not to be defined in terms of the distribution of functionings. And to define "so and so much" in terms of the distribution of functionings would again generate the problem of common diseases.

In contrast to Hausman's definition of 'efficiency', the DPE-theory's definitions of survival-efficiency and reproduction-efficiency allow us to order the efficiencies of different SDPs not only on an ordinal scale, but on a ratio scale. This makes it possible for the DPE-theory to distinguish survival-health

from survival-pathology (or, respectively reproduction-health from reproduction-pathology) by a general line indicating what is significantly worse than the standard.

Let us discuss in more concrete terms how the DPE-theory serves better than the biostatistical theory with regard to common diseases. The DPE-theory serves better than the BST, since it allows for very large proportions of the tokens of a trait type being survival-pathological or reproduction-pathological. For example, if dental caries reduces the survival-efficiency of the dental system's SDP (specific disposition profile) for decomposing food below k_s, and occurs among, say, 20 percent of the reference class, then tooth decay is a survival-pathological condition in 20 percent of the reference class members. The same goes for canine hip dysplasia in dogs. If it reduces the survival-efficiency of the hip's SDP for some of the hip's physiological functions (perhaps stabilizing the body while walking) below k_s, and this state occurs among 30 percent of the dogs in a certain breed, then it is a survival-pathological state among these 30 percent.

The DPE-theory may even ascribe survival-pathology or reproduction-pathology to 100 percent of the trait tokens. Although such a situation is very unlikely, it is at least theoretically possible. Consider the following example about teeth and cavities, originally from Hausman (2012, 536). Let us assume that it is readily attainable for a significant share of a human reference class, in common environments of this reference class, to have no cavities. Assume also that teeth without cavities are the teeth that are best disposed for a high further life expectancy. This means that the survival-exemplary CDP (complete disposition profile) of the concerned reference class has SDPs for the teeth's physiological functions that represent teeth with no cavities. However, if all reference class members, despite this, choose to eat a lot of sugar they may all obtain cavities and, thereby, survival-pathological teeth.

6.2 Solving Kingma's Dilemma

Let us now consider Kingma's dilemma. We saw that Kingma (2010) argues that a theory like the biostatistical theory cannot account both for the situation-specificity of physiological functions and situation-specific pathology. In order to account for the situation-specificity of physiological functions, what is a "normal performance" must be considered relative to the situation of the

performance.[28] For example, a healthy human heart will beat with a lower beating frequency when the individual is resting than when the individual is running. However, Kingma argues, if one considers the normalcy of a performances relative to the situation in which it is performed, it becomes impossible to account for situation-specific diseases. According to Kingma, if a necessary condition for pathology is abnormal performance, then situation-specific diseases cannot count as pathological since the performances of physiological functions in cases of such diseases are normal in the situations in which they occur.

In Chapter 2 (Section 2.3.3), I argued in line with Hausman (2011) and Boorse (2014) that Kingma's dilemma is no genuine dilemma, and hence that it is not a problem for the BST. Basically, Kingma's mistake is that she does not distinguish between how tokens of a trait type are normally affected by situations and how trait tokens are normally disposed to respond to different situations.

However, Kingma's alleged dilemma has been given much attention, and is still often considered to be a genuine dilemma. The disposition profile efficiency theory, I think, has the resources to explain, better than Hausman's and Boorse's respective theories, that the alleged dilemma is not a threat for theoretically motivated theories of health and pathology. Beneath I will use the DPE-theory to show that Kingma's dilemma is no genuine dilemma (Section 6.2.4). First, though, I will discuss earlier suggested solutions, besides that of Hausman (2011) and Boorse (2014). These are suggestions by Dussault and Gagné-Julien (2015) (Section 6.2.1), Kraemer (2013) (Section 6.2.2), and Garson and Piccinini (2014) (Section 6.2.3).

6.2.1 Appeal to Biological Normality

Dussault and Gagné-Julien (2015) claim to solve Kingma's dilemma by appealing to a non-statistical idea of design. In their rather complex account of health and pathology, they suggest the following: "An organism is healthy if and only if it is intrinsically disposed to homeostatically maintain or restore its intrinsic disposition to perform its designed functions in relevant situations" (Dussault and Gagné-Julien 2015, 75). Otherwise there is pathology in

[28] For the biostatistical theory, a "normal performance" means a statistically typical performance for the reference class members. For the disposition profile efficiency theory, though, a "normal performance" is not purely statistically determined, but should roughly be understood as the most beneficial performance that is readily attainable for a significant share of the reference class members in their normal environments.

the organism. The idea of "designed functions" they appeal to here is non-statistical. So, in contrast to Boorse, what sets the standard that serves as comparison in determinations of health statuses is not what is a statistically typical performance in the reference class, given the situation. How, then, is the standard set? To explain the standard, Dussault and Gagné-Julien refer to a concept of 'biological normality' discussed by Wachbroit (1994). Wachbroit argues that there is a special concept of biological normality which function statements in biology aim at. And this concept, he argues, is not statistical. Although Wachbroit provides arguments for the claim that the concept of biological normality is not statistical, he does not give any positive account of biological normality. And if we turn to Dussault and Gagné-Julien (2015), they do not say anything substantive about the concept of design they appeal to, other than that it is not statistical.

Although I agree with Dussault and Gagné-Julien that the standard for health should not be determined by what is typical in a reference class, I am hesitant toward their reference to biological normality. As long as the term 'biological normality' is not given some clearer meaning than being distinct from statistical normality, it is hard to evaluate Dussault and Gagné-Julien's definitions of 'health' and 'pathology'.

6.2.2 Allow for Differently Fine-Grained Descriptions of Situations

Kraemer's (2013) idea for solving Kingma's dilemma is to allow for a range of specifications of circumstances, differing in granularity of description. Kraemer uses the example of a person who overdoses on paracetamol, which (as a typical effect of overdosing on paracetamol) results in liver failure. In this case, Kraemer points out, the situation may be described as comprising the fact that the person *has consumed an overdose* of paracetamol. But the situation may also be given an alternative description, at a more course-grained level of granularity, as comprising the fact that the person *has consumed* paracetamol. With the first description the liver failure is typical given the situation as described, whereas with the second, less granular, description the liver failure is not typical given the situation as described. Hence, according to Kraemer, we may use the less granular description to account for the liver failure as pathological.

I think that Kraemer's idea for solving Kingma's dilemma is problematic. For the idea to handle all cases of situation-specific pathology, we will sometimes need to go absurdly coarse-grained. In Kraemer's example, where the

difference is between consuming an overdose of paracetamol and consuming paracetamol, the shift in granularity may perhaps not sound too strange. However, in other cases we would need to go considerably less fine-grained in order to get the results we want. Consider, for example, a case where a person eats a poison that severely harms her digestive system. Let us assume that ingesting this poison in any amount leads to pathology. In order to get the result that the digestive system is pathological, we would therefore need to describe the situation as coarse-grained as having ingested something (rather than having ingested poison). Or consider a person stepping on a landmine.[29] The normal reaction to stepping on a landmine is to have at least one's leg injured. In order to get the result that the injured leg is pathological, we would have to make the description as coarse-grained as having taken a step.

One may, however, object that my example cases do not show that there is anything wrong with Kraemer's suggestion. Perhaps the best result we can get when trying to define theoretical concepts of health and pathology amounts to health and pathology being arbitrarily relative to our descriptions of situations. However, I think that it is possible to define 'health' and 'pathology' in a way that is not arbitrarily relative to how we describe situations. In the disposition profile efficiency theory, situations are states of the world, not descriptions (of some granularity) of such states. And, if it is possible to define 'health' and 'pathology' as not arbitrarily relative to how we describe situations, as the DPE-theory shows that it is, that should be the preferable option.

6.2.3 Consider the Trait Token's Disposition to Perform the Function in Different Situations

I presented Garson and Piccinini's "biostatistical account of functions" above when discussing the problem of common diseases (Section 6.14). Garson and Piccinini (2014) take their account to solve not only the problem of common diseases, but also Kingma's dilemma. Their idea for solving Kingma's dilemma is by considering how the trait token would perform its physiological functions in different situations than that actually present.

In order to show how their account solves Kingma's dilemma, Garson and Piccinini apply it to Kingma's own example cases, for instance about Carol and her digestive system. Let us briefly recap these cases. In the first case, Carol is resting after a meal and her digestive system is fully active. Here the

[29] This example is borrowed from Hausman (2011), although he uses it to make a different point.

intuitively correct judgement is that the digestive system is healthy. In the second case, Carol has been fasting, and her digestive system is inactive. Again, the intuitively correct judgement is that the digestive system is healthy. In the third case, Carol is in the middle of a heavy exercise session, and her digestive system is rather inactive. Once again, the intuitively correct judgement is that the digestive system is healthy. In the fourth case, Carol has ingested a poison that has paralyzed her digestive system, and hence the digestive system is inactive. Here the intuitively correct judgement is that the digestive system is pathological. Garson and Piccinini reach the following conclusions about these cases:

> In *S1* (the individual is resting after a meal), the situation is appropriate for digestion and digestion should occur at a relatively high rate. By hypothesis, digestion is occurring, so the stomach is not dysfunctional. In *S2* (the individual has been fasting), the digestive system is not digesting food. This is as it should be because […] if the organism is in *s* ($s \notin S$), the rate of functioning of zero (or close to zero) is appropriate. Furthermore, by hypothesis, the stomach is still able to digest should the situation become appropriate (for example, by the ingestion of food). Thus, the stomach is not dysfunctional. Situation *S3* (the organism has been exercising heavily and hence is digesting at a low rate) is similar to *S2*, except that the situation is such that the appropriate rate of functioning is now low rather than nil.
>
> In contrast, in *S4* (the individual has ingested a poison which paralyzes the digestive tract), the stomach is dysfunctional because it cannot digest food. That is, there is at least one *s* ($s \in S$) such that even if the individual were placed in *s*, the digestive system would not digest food. Because of this, the stomach is dysfunctional. (2014, 16)

Garson and Piccinini are, I think, correct in their conclusions. As also Hausman and Boorse point out, as we saw in Chapter 2 (Section 2.3.3), Garson and Piccinini press that the relevant comparison standard is not relative to the particular situation. Rather, the relevant comparison is how trait tokens of the type in question are generally disposed to behave in different situations. In principle, Garson and Piccinini solve Kingma's dilemma. However, as I pointed out above in Section 6.1.4, their account only defines 'dysfunction', and not 'pathology'. In contrast to Garson and Piccinini's biostatistical account of function, the disposition profile efficiency theory is able to compare how well non-perfectly functioning trait tokens function. Thereby it also becomes possible for the DPE-theory to define 'pathology' in terms of a sufficiently reduced functioning, and not just in terms of non-perfect functioning. Let us next discuss how the DPE-theory solves Kingma's dilemma.

6.2.4 The Disposition Profile Efficiency Theory

Already in Chapter 2 (Section 2.3.3), I argued, with reference to Hausman's (2011) and Boorse's (2014) arguments, that Kingma's dilemma is not a genuine dilemma for the biostatistical theory. We have now also seen by Garson and Piccinini's biostatistical account of functions that the dilemma is, in principle, possible to solve. The general conclusion is that Kingma's dilemma is based on a confusion between the question whether it is normal for a token of a trait type to acquire some disposition in some situation S and the question whether the disposition it acquires in S is normal.

The disposition profile efficiency theory captures this conclusion particularly well. A trait token is survival-healthy if and only if all of its SDPs (specific disposition profiles) are such that, when considered as alternative parts of a survival-exemplary CDP (complete disposition profile), they do not give rise to a much lower further life expectancy compared to the original survival-exemplary CDP. And a trait token is survival-pathological if and only if some of its SDPs is such that, when viewed as an alternative part of a survival-exemplary CDP, it gives rise to a significantly lower further life expectancy compared to the original survival-exemplary CDP. In these definitions of survival-health and survival-pathology there is no reference made to the present situation of the trait token. Whether a trait token is survival-healthy or survival-pathological only depends on how well its SDPs generally (i.e. not dependent on any particular situation of the individual) serves the goal of survival.

To see how the DPE-theory deals with Kingma's dilemma, let us apply the theory to Kingma's four test cases of Carol's digestive system. In these examples, I will assume that the digestive system has one and only one physiological function – to digest. I will also assume that Carol belongs to exactly one reference class and that there is, for this reference class, exactly one survival-exemplary CDP. Let us also suppose that Carol is 30 years old.

Case 1: Carol relaxes after a meal

Consider Kingma's case of Carol's digestive system when Carol is relaxing after a meal. In order to determine the survival-health status of Carol's digestive system, we must consider the survival-efficiencies of its SDPs. In this example, this means that we need to determine the survival-efficiency of Carol's digestive system's SDP for digesting. I will denote the survival-efficiency of Carol's digestive system's SDP for digesting '$s\text{-}eff$(Carol's digestive system's SDP)'.

Let us make some concrete assumptions about further life expectancies for this example. First, suppose that the further life expectancy for the survival-exemplary CDP of Carol's reference class is 50 years. What is then the further life expectancy of the survival-exemplary CDP manipulated with Carol's digestive system's SDP for digesting? Since there is nothing special about Carol in this case, we may assume that her digestive system is like the digestive system of the survival-exemplary CDP. This means that the further life expectancy of the survival-exemplary CDP manipulated with Carol's digestive system's SDP for digesting should be the same as the further life expectancy for the survival-exemplary CDP, i.e. 50 years. The survival-efficiency that we get is:

$$\text{s-}eff(\text{Carol's digestive system's SDP}) = \frac{50}{50} = 1$$

This survival-efficiency is above k_s. We may hence conclude that Carol's digestive system is survival-healthy.

Case 2: Carol has fasted

Consider the case of Carol's digestive system when Carol has fasted for a period. Here, the survival-health status evaluation will be identical to the previous one, since Carol's SDP for digestion is exactly the same. It does not matter that the situation is different, since the survival-efficiency of the digestive system is not relative to the present situation (more than it is to any other potential situation for a digestive system). Hence, the survival-efficiency is again 1, which means that Carol's digestive system is survival-healthy.

Case 3: Carol is exercising

Consider the case of Carol's digestive system when Carol is in the middle of a lengthy exercise session. Again, Carol's SDP for digestion is exactly the same as in the previous cases, which means that the survival-health status evaluation will be identical to the one in the first and second cases. Hence, the survival-efficiency is 1 once more, which means that Carol's digestive system is survival-healthy.

Case 4: Carol has ingested poison

Consider the case of Carol's digestive system when Carol has ingested poison. Suppose that the poison affects Carol's digestive system so that it no longer digests properly. The newly acquired disposition is represented by an SDP that reports of no digestion in any situation. Here it is clear that Carol's SDP

for digestion is different from the survival-exemplary CDP's SDP for digestion. Assume that the further life expectancy of the survival-exemplary CDP, manipulated with regard to Carol's digestive system's SDP for digestion, is 0,08 years (roughly one month). We then get:

$$\text{s-}eff(\text{Carol's digestive system's SDP}) = \frac{0{,}08}{50} = 0{,}0016$$

0,0016 should be well below k_s. Hence Carol's digestive system is survival-pathological. We may thereby conclude that the DPE-theory has no problem of handling Kingma's cases about Carol.

Even if Kingma's dilemma is, as I have argued, in principle no problem for the BST, the BST faces a related problem, which I mentioned in relation to the discussion of Kingma's dilemma in Chapter 2 (Section 2.3.3). The problem is that in the BST's definitions of 'health' and 'pathology', Boorse only takes into consideration how the trait token is disposed to perform its type's physiological functions on *typical occasions* (i.e. occasions in which the physiological function is typically performed). This gives, as I explained, the strange result that a trait token's disposition to carry out a physiological function at full activity in a situation in which it is not adequate to carry out the function at all does not make the trait token pathological, even if this disposition significantly lowers further life expectancy. Examples where a physiological function is carried out in situations, in which it should not, are cases of autoimmune diseases, where the immune system is highly active although there is no threat to incapacitate. Since the DPE-theory takes into consideration not only potential performances in typical situations, but in all situations, it does not have a problem to account for the survival-pathology of an immune system highly active in situations where there is no threat to incapacitate. This is because in the survival-exemplary CDP's SDP for incapacitating cells or other bodies, the feature value configurations report a fairly low activity for situations where there is no threat to incapacitate. If the feature value configurations for those situations were substituted with feature value configurations instead reporting a very high activity, the further life expectancy would decrease.

Let us sum up. As already indicated by Hausman (2011), Boorse (2014) and Garson and Piccinini (2014), Kingma's dilemma is no genuine dilemma. By carefully modelling dispositions for physiological functions and defining not only 'dysfunction' but also 'pathology', the DPE-theory shows more

clearly than Hausman, Boorse, and Garson and Piccinini that Kingma's dilemma is no problem for purely theoretically motivated theories of health and pathology.

7 Contested Cases

I mentioned in the introductory chapter that there are some types of conditions, for which it is contested whether they should count as healthy or pathological. In this chapter, I will discuss these types of conditions in relation to the disposition profile efficiency theory. I will start with conditions that are trivial from the perspective of the whole organism (Section 7.1). Then, I will discuss conditions of normal ageing (Section 7.2), defense mechanisms (Section 7.3), and, lastly, risk conditions (Section 7.4).

7.1 Trivial Conditions

There is disagreement about whether conditions where a trait token is not functioning "properly", but where this does not have any gross effect on the organism as a whole (it does, for example, not reduce the organism's further life expectancy or its further expected number of offspring), should count as healthy or pathological. For example, whereas Boorse (2014, 706) thinks that a single malfunctioning cell should count as pathological, Nordenfelt (1987, 28) thinks that it should not.

Let us consider what the disposition profile efficiency theory implies about this issue. Since the DPE-theory defines 'survival-pathology' as a significantly reduced survival-efficiency, where a significantly reduced survival-efficiency is spelled out as a significantly lower further life expectancy than that of the standard, the DPE-theory judges conditions that do not significantly reduce the further life expectancy of the whole CDP (complete disposition profile) healthy. This means that conditions that are trivial for the organism's further life expectancy are not survival-pathological on the DPE-theory. It even means that it might be impossible for tokens of some trait types to become survival-pathological. For tokens of trait types whose physiological functions contribute only little or marginally to further life expectancy, it becomes very unlikely, or even impossible, that their SDPs (specific disposition profiles) have a survival-efficiency at or below k_s. Mutatis mutandis for reproduction-pathology.

Is this problematic? I will argue that although it is in a sense somewhat of a cost for the theory, this cost is not as large as one may think, and it is a cost that is worth paying. Below, I will discuss two types of cases where it will be impossible, or very unlikely, for a token of a certain trait type to count as survival-pathological.

Note that even if very small survival-efficiency reductions cannot by themselves constitute survival-pathology, they may still *contribute to* survival-pathology. Recall that organisms comprise traits and physiological functions at different levels. For example, in comparison to the circulatory system, the heart is a trait at a lower level, and in comparison to the heart, a heart cell is a trait at a lower level. A small survival-efficiency reduction at a low level in an organism may contribute to survival-pathology at a higher level in that organism. Consider, for example, type-1 diabetes mellitus, where the beta cells in the pancreas, which are responsible for insulin production, get destructed. As mentioned earlier, type-1 diabetes mellitus is a serious disease with symptoms such as fatigue, nausea, blurry vision, and weight loss. What is noteworthy here is that what causes the detrimental states is merely a dysfunction of certain cells. Since there are so many beta cells, a single beta cell's SDP for insulin production will have a survival-efficiency of 1 (or very close to 1) irrespectively of whether it is able or unable to produce insulin. This means that, in cases of type-1 diabetes mellitus, the DPE-theory cannot say of a single defective beta cell that it is survival-pathological. This might sound like a problematic result. However, I think it is not. That we cannot say of each beta cell that it is survival-pathological does not rule out that there is survival-pathology in the organism. In cases of type-1 diabetes mellitus, there are larger survival-efficiency reductions at higher levels than at the level of the single cell. Considering the whole system of beta cells, the survival-efficiency of its SDP for producing insulin is going to be well below the value of k_s. Thus, when an insulin-producing system breaks down, as it does by type-1 diabetes mellitus, its SDP's survival-efficiency for producing insulin becomes significantly reduced, and thus survival-pathological. Moreover, several effects of the insulin deficit on tokens of other trait types in the organism may lead to further survival-pathological states.

If we instead consider physiological functions that only marginally contribute to survival or reproduction, and which are not part of a higher-level system, we may find trait tokens that cannot in any sense become pathological, or, at least, are very unlikely to become pathological. Consider, for example, the stapedius reflex – a muscle contraction in the middle ear in response to loud sounds, which decreases the transmission of vibrational energy to the

cochlea. This function contributes to survival by protecting the cochlea from damage caused by excessive stimulation, which in turn protects the hearing system's capacity for distinguishing sounds. However, if we compare the effects on further life expectancy by a diminished stapedius reflex with a diminished capacity of the heart to pump blood, the difference is quite large. Even if the value of k_s is high enough to allow for the muscle responsible for the stapedius reflex to become survival-pathological, this will require larger functional deviations from the survival-exemplary CDP than those required for a heart token to count as survival-pathological.

Although this may to some extent sound counterintuitive, I think it can be explained. In medical theory, small deviations from the standard heart are often regarded as rather serious, whereas small deviations from the standards for trait types that are not as important for further life expectancy (like the stapedius reflex) are not. It should also be noted that, even if there are tokens of trait types that cannot become survival- or reproduction-pathological, we may still make meaningful claims about the survival- or reproduction-efficiencies of these token's SDPs'. Although these efficiencies cannot not vary enough for a trait token to have such a reduced efficiency that it counts as pathological, they may still vary on a very detailed scale.

I have now concluded, first, that very limited survival-efficiency reductions at lower levels, although not by themselves enough to constitute pathology, may contribute to pathology on higher levels, and, second, that it is not such a strange result of the DPE-theory that it is unlikely, or even impossible, for tokens of some trait types to become pathological. To the extent that the discussed implications are still counterintuitive, I think that this price is worth paying for an explicative account of health and pathology. In order to avoid the implication that it will require more of, or be impossible for, tokens of some trait types to count as survival-pathological, we would either have to allow for different distinguishing lines between survival-health and survival-pathology for different trait types, or we would have to allow for further relevant goals besides survival and reproduction. None of these alternatives are promising.

If we were to allow for different distinguishing lines between health and pathology for different trait types, the account would become arbitrary – there would be no justifiable basis for the distinguishing line. Rather, it seems, the lines separating health from pathology for different trait types would be based only on our pre-theoretical judgements of which states are healthy and which are pathological. This means that the theory would not provide us with much understanding of health and pathology as general concepts. Here, we should

choose fruitfulness and consistency over conformity to language usage and hence hold on to the idea that there is a general line distinguishing survival-pathology from survival-health, and a general line distinguishing reproduction-pathology from reproduction-health.

The idea to allow for further goals besides survival and reproduction was discussed in Chapter 3 (Section 3.1.1), where I concluded that it seems doubtful that there is some further goal, besides survival and reproduction, that is relevant for physiological functions, health, and pathology. As was concluded when discussing different suggestions for further goals, it seems that adding further goals besides survival and reproduction will make the account deviate even more from language usage in medical theory.

7.2 Normal Ageing

It is contested whether conditions of normal ageing should count as healthy or pathological. For example, Boorse (1977, 567) acknowledges such conditions as healthy since they are typical in reference classes of high age, but later (Boorse 2014, 714) discusses the possibility of judging such conditions pathological by treating age as irrelevant after adulthood. Brown (1985, 317) and Pawelzik (1990, 18) argue that conditions of normal ageing should be considered pathological because they are subject to interventions in medical practice.[30]

To clarify the discussion of normal ageing, I suggest we distinguish between two types of states that may both be aimed at by the term 'normal ageing'. States of the first kind are functional impairments that inevitably occur among more or less all individuals while getting older. Two examples here are reduction in muscle function due to reduction in muscle mass and blurry vision at short distances due to hardening of the lens (presbyopia). States of the second kind are conditions that merely become more common in groups of individuals as they get older. States that could serve as examples here are Schwartz's (2007) cases of urinary dysfunction due to benign prostatic hypertrophy among men older than 70, and senility of the Alzheimer type among people over 85. These states may be described as a subcategory of common diseases. Although there is no sharp line between the two types of normal ageing, I think the distinction is useful. While states of the second kind are

[30] Brown (1985, 317) and Pawelzik (1990, 18) also present this view on conditions of normal ageing as an argument against Boorse's theory.

thought of as pathological in medical theory, states of the first kind are not (at least it is controversial to claim that they are).

If we consider the biostatistical theory, it classifies both states of the first kind, and (at least most) states of the second kind, as healthy – simply because they are both rather common. This result has given rise to criticism, to which Boorse (1997, 91-92) has answered by stating that there is no contradiction between a state being healthy and health care aiming to treat it. Yet, while this is correct, there may be other reasons for thinking that some of these states should count as pathological. Considering pathologists' classification of states into healthy and pathological, it is a cost for the BST that it deviates from medical theory to a quite large extent at this issue.

What does the disposition profile efficiency theory say about states of normal ageing? Let us first consider states that inevitably affect more or less all individuals while getting older. On the DPE-theory, just like the BST, these conditions are survival-healthy: As long as these states are the best readily attainable states for reference classes of higher ages, they will be instantiated in the survival-exemplary CDP of these reference classes. This means that an SDP (specific disposition profile) affected by the type of normal ageing that inevitably affects more or less all individuals while getting older will typically not have a survival-efficiency below k_s, and that trait tokens affected by this sort of normal ageing are typically survival-healthy.

For illustration, consider the human individual Hedda who is 74 years old. We are interested in evaluating the survival-health status of Hedda's eyes. Hedda has presbyopia, but no other problems with her sight. Presbyopia is, as indicated above, a condition that affects most humans as they get older. This is because the lens gets stiffer when it gets older, which in turn makes it hard for the eye to focus on close objects.

In order to evaluate the survival-health status of Hedda's eyes, we must first evaluate the survival-efficiency of Hedda's eyes' SDPs for the physiological functions of eyes. Let us suppose that the eye has one and only one physiological function, which is to detect optical information. For the purpose of this example, let us also assume that Hedda belongs to one unique reference class at the time of the evaluation and that there is one unique survival-exemplary CDP (complete disposition profile) for this reference class. Given these assumptions, let us denote the survival-efficiency of Hedda's eyes' SDP for detecting optical information '$s\text{-}eff(\text{Hedda's eyes' SDP})$'.

Suppose that the further life expectancy for the survival-exemplary CDP of Hedda's reference class is 10 years. Given this, what should we expect the further life expectancy of the survival-exemplary CDP manipulated with

Hedda's SDP for detecting optical information to be? Since lenses inevitably stiffen with age, it is not readily attainable for a significant share of Hedda's reference class members to have eyes able to focus on close objects. Hence, Hedda's eyes' SDP for detecting optical information does not differ from the corresponding SDP in the survival-exemplary CDP. And, hence, the further life expectancy of the survival-exemplary CDP manipulated with Hedda's eyes' SDPs for detecting optical information is the same as that of the survival-exemplary CDP, i.e. 10 years. Assuming these numbers we get:

$$s\text{-}eff(\text{Hedda's eyes' SDP}) = \frac{10}{10} = 1$$

This survival-efficiency is above k_s. From this we may conclude that Hedda's eyes are survival-healthy.

However, if we consider states that merely become more common among individuals as they get older, there is a difference. In these cases, the survival-exemplary CDP's further life expectancy more or less remains unaffected by the condition in question. So, here the DPE-theory gives the same results as it does in cases of common diseases. It ascribes survival-pathology to more states than the BST.

For illustration, consider the human individual Isabel who is 90 years old. We are interested in evaluating the survival-efficiency of her brain. Isabel has Alzheimer's disease, which is rather common among older people. As reported by Schwartz (2007, 375), it affects 16 percent of people older than 85. In cases of Alzheimer's disease, there is a loss of neurons and synapses in the brain, which results in atrophy of affected regions, typically including regions of the temporal lobe, the parietal lobe, the frontal cortex and the cingulate gyrus. This process affects several brain functions, concerning for example learning and memory, language, perception, and execution of movements. Although the disease does usually not cause death by itself, it often indirectly causes death by unabling the individual to cope with infections, for example of pressure ulcers or pneumonia.

In order to evaluate the survival-health status of Isabel's brain, we must first evaluate the survival-efficiency of her brain's SDPs for the physiological functions of brains. The brain is associated with numerous physiological functions. In order for it to be survival-pathological on the DPE-theory, it is enough that its SDP for one of these physiological functions has a significantly reduced survival-efficiency (for all choices of reference class and survival-

exemplary CDP). Let us here focus only on the physiological function of remembering. Let us for the sake of simplicity assume in this example that Isabel belongs to one unique reference class at the time of the evaluation and that there is one unique survival-exemplary CDP for this reference class. Given these assumptions I will denote the survival-efficiency of Isabel's brain's SDP for remembering '$s\text{-}eff$(Isabel's brain's SDP)'.

Suppose that the further life expectancy of the survival-exemplary CDP for Isabel's reference class is 4 years. For most elderly humans, it is not readily attainable to remember for example as well as most 20-year-old humans. However, although Alzheimer's disease and other types of dementia affect quite many individuals in Isabel's reference class, most individuals in the reference class are not seriously affected by such conditions. It is still readily attainable for a significant share of the reference class members to remember better than Isabel. Since I think it is reasonable to assume that memory is to some extent important for high further life expectancy (given the exclusion of specially directed medical and social interventions), the survival-exemplary CDP manipulated with Isabel's brain's SDP for remembering will have a lower further life expectancy than the survival-exemplary CDP. Suppose that the further life expectancy of the survival-exemplary CDP manipulated with Isabel's brain's SDP for remembering is one year. From these numbers we get:

$$s\text{-}eff(\text{Isabel's brain's SDP}) = \frac{1}{4} = 0{,}25$$

This survival-efficiency is rather low. Given that k_s should be located closer to 1 than 0, a survival-efficiency of 0,25 reasonably indicates survival-pathology. We may hence conclude that Isabel's brain is survival-pathological.

7.3 Defense Mechanisms

A type of condition, where it is perhaps not obvious at first sight whether the condition is healthy or pathological concerns defense mechanisms. In discussions of the biostatistical theory, Nordenfelt (1987; 2001) and Worrall and Worrall (2001) express different views on whether the activation of defense

mechanisms such as fever, vomiting, and diarrhea are healthy or pathological.[31]

According to Nordenfelt, activated defense mechanisms like the ones just mentioned should count as pathological. He writes:

> We shall, in particular, study the situation where a number of microbes invade a tissue, i.e. the paradigm case of an infection.
>
> Consider the main steps in such an infectious process. A number of pathogenic agents enter, say, the throat. They start producing toxins, which immediately destroy a great number of cells on the mucous membranes. The body quickly reacts to this attack. There is a great concentration of blood at the focal points of the infection; the body temperature rises; certain tissues create antibodies against the viruses and the pathogenic toxin. As a result the toxin is gradually neutralized and the microbes killed.
>
> This outline of a description of a process of infection is at the same time a description of a species-typical contribution to the ultimate goals of survival and reproduction. In fact, the infectious disease can be seen as the species-typical reaction to the circumstance of a certain microbial invasion.
>
> But this, then, becomes paradoxical. A typical disease can be seen, on the BST, as a species-typical reaction, i.e. as a healthy response to a difficult environment. (Nordenfelt 1987, 30)

In a later text, Nordenfelt again discusses the same example:

> To summarize: my point with the infection example was the following. Assume that a pathogen is so virulent as to trigger a strong but still statistically normal response, given those circumstances which include the pathogens. The strong response effected by the antibodies may lead to significant bodily change and a change in functional activity. The properties of this change, in terms of inflammation, high temperature and, on the illness side, pain, fatigue, and disability, are quite well-known. We then have a typical disease (as recognized in medical classifications) which is constituted by a normal bodily response, in the sense of making a statistically normal contribution to survival, in relation to a harsh environment. (2001, 17)

[31] Nordenfelt (1987; 2001) and Worrall and Worrall (2001) also interpret the BST differently. Whereas Nordenfelt thinks that the BST classifies activated defense mechanisms as healthy, Worrall and Worrall in contrast think that the BST does not account for the activation of defense mechanisms as healthy. Here I think that Nordenfelt's conclusion is correct. Since the activation of defense mechanisms such as fever and vomiting in an organism who is infected by a virus, or who has eaten something poisonous, typically contributes to the organism's survival, these defense mechanisms should count as physiological functions on the BST. And given that a specific individual is infected by a virus, or has been eating something poisonous, fever or vomiting are just, typical, healthy responses.

Hence, according to Nordenfelt, the activation of defense mechanisms should count as pathological, since they are typical components of diseases (e.g. infections).

In contrast to Nordenfelt, Worrall and Worrall argue that the activation of defense mechanisms such as the ones mentioned above are healthy. They write:

> […] as recent Darwinian approaches to medicine have emphasized, fever may very well be an adaptive *beneficial* response to infection – beneficial because it does more damage to the invading bacteria or viruses than to the host. Similarly diarrhoea and vomiting, far from implying failure of some function, may well in fact be examples of adaptive functioning – one way in which the host organism can efficiently expel large numbers of the invaders. […] These Darwinian insights, by the way, underline the lesson (already absorbed in the more enlightened medical quarters) that the standard medical approach to treating fever, diarrhoea and vomiting may sometimes be entirely misguided – prolonging the condition or endangering the patient, rather than alleviating the condition and helping the patient. (Worrall and Worrall 2001, 44-45, italics are original)

So, we have on the one hand arguments for conditions such as fever, vomiting, and diarrhea being pathological, and on the other hand arguments for them being healthy. Although more or less unreflective intuitions about these conditions may vary, I think there is a determinate answer to the question whether the activation of defense mechanisms should be considered healthy or pathological. And this answer is very much in line with the view of Worrall and Worrall: the activation of defense mechanisms are typically healthy.

As also Nordenfelt (1987) notes, in situations where there is a microbial threat, typical responses such as fever and production of antibodies are beneficial for survival. Although it would not be beneficial for high further life expectancy to have fever, or diarrhea, or to vomit in situations without microbial threats, it is beneficial to have these reactions in situations of microbial threats. This is something that is well illustrated by the disposition profile efficiency theory. In a trait token's SDP (specific disposition profile), there is for each situation a feature value configuration reporting of the trait token's activity. The survival-exemplary CDP (complete disposition profile) will have the combination of SDPs that, among those that are readily attainable for a significant share of the reference class members in their common environments, gives the highest further life expectancy. For example, with regard to the physiological function of temperature regulation, the survival-exemplary

CDP of some human reference class will have feature value configurations reporting of higher temperatures in situations of microbial threats and lower temperatures in other situations. This is the standard for survival-health. What would be pathological for a trait token regulating body temperature is not to increase the temperature in situations of microbial threats, but to *not* increase the temperature in such situations. It would also be survival-pathological to increase the temperature so much that the reaction instead lowers further life expectancy, or to increase the temperature also in cases where there is no microbial threat.

Nordenfelt is right that when there is, for example, fever in the organism, there is often also pathology in the organism. However, what is important to clarify here is which trait token (or tokens) that is (are) pathological. In Nordenfelt's own example cited above, it is not the trait tokens responding to the microbes by defense mechanisms that are pathological. Rather, it is the trait tokens attacked by the microbes (e.g. the mucous membranes) that are pathological. Hence, although activated defense mechanisms indicate pathology in the organism, the trait tokens responsible for these reactions are not themselves pathological. To have an infection, but no response by the immune system, on the other hand, would indicate that the immune system is pathological.

A further mischaracterization in Nordenfelt's description is, I think, that pathological states are not clearly distinguished from diagnostic criteria. When diagnosing an infection, it is typical to consider whether the patient has fever. However, as pointed out above, fever is only indicating that something in the organism is pathological; it is not itself pathological. Diagnostic criteria and pathology must be separated.[32]

7.4 Risk Conditions

The last type of debatable condition that I will bring up to discussion are risk conditions. Examples mentioned in the literature (e.g. Boorse 2012; Schwartz 2008) are hypertension, high cholesterol, osteoporosis, and obesity. The issue of risk conditions is complex, and I will not be able to discuss it exhaustively here. However, since I think that the disposition profile efficiency theory has resources to enrich the debate about these sorts of conditions, I want to bring

[32] See also Boorse's response to Nordenfelt (Boorse 1997, 85; 2014, 712-713).

the issue to light and at least sketch some arguments and suggestions for further research.

Schwartz (2008) points out that, in clinical practice, risk conditions (or "modifiable risk") are often regarded as pathological since treating such states may prevent future non-beneficial states. For example, Schwartz writes:

> What matters is that for patients with blood pressure in this range [a systolic blood pressure of 140-159 mm Hg or a diastolic blood pressure of 90-99 mm Hg], there is good evidence that using available medications to lower the blood pressure also reduces cardiovascular risk. As the guidelines state: "The ultimate public health goal of antihypertensive therapy is to reduce cardiovascular and renal morbidity and mortality" […]. (2008, 321-322)

But, according to Schwartz, risk conditions and pathology are basically different things, and should therefore be clearly separated.[33] Schwartz writes:

> First, there is the simple point that we should call these conditions what they are. The growing attention paid to identifying and reducing risk should not be allowed to warp the notion of disease. It would be regrettable if the desire to improve preventive care resulted in an unnecessary blurring of the distinction between health and disease. (2008, 332)

According to Schwartz (2008, 325-329), risk conditions cannot constitute pathology because pathology requires present dysfunction. Conditions like stage 1-hypertension and high cholesterol do usually not cause any present dysfunction, Schwartz argues: there is nothing in the organism that, because of the condition, does not work adequately. Stage 1-hypertension and high cholesterol merely cause an increased risk of future dysfunction. Hence, according to Schwartz, these states are merely conditions of risk, not of pathology.

When Schwartz claims that risk cannot constitute pathology because pathology requires present dysfunction, he uses the biostatistical theory as an example of a dysfunction-requiring account (2008, 325). However, Boorse (2012) does not fully agree with Schwartz's conclusions. Boorse (2012, 19-20) mentions physiological functions where it is clear that their benefit for survival or reproduction is only associated with future situations. For example, embryological origins of the eyes do not make a difference for the individual

[33] Schwartz (2008) also points out that, as long as one does not limit the set of conditions that should entitle a person to public health care to states of pathology, there is no worry in not classifying risk conditions as pathological.

at the time, but surely for the ability to see later. And some physiological functions are specifically disease preventing, for example the lymphocytes' physiological function of preventing infectious disease. That these physiological functions are not important for surviving at the moment, but only for surviving in the future, does not mean that embryological eyes or lymphocytes (in the absence of infection) cannot at the moment be pathological. According to Boorse, if a state impedes survival or reproduction, then there seems to be a reduction of efficiency.

If we consider the DPE-theory, it is also apparent that future aspects matter for health status evaluations. First, health and pathology are about dispositions to perform physiological functions in different situations, rather than about mere actual performances. Second, what determines the health status of a trait token are the efficiencies of its dispositions for performing physiological functions. These efficiencies are clearly future-oriented, since they concern further life expectancy or further expected number of offspring. It is hence questionable whether it is, as Schwartz thinks, conceptually inadequate to include considerations of risk in health status determinations.

However, in order to fully understand risk conditions in relation to pathology, more work needs to be done. Let me sketch some issues that are worth investigating.

First, it needs to be made clear what risk conditions are. A very basic question which needs to be dealt with here is what the risk is a risk *of*. Is it, for example, risk of pathology, or risk of having a reduced length of life left, or something else? Also, it will be useful to consider, as Boorse (2012, 14) does, a distinction between "causal risk factors" (factors that may causally contribute to the thing that there is risk of) and "risk markers" (factors that merely correlate with the thing that there is risk of). According to Boorse, although both sorts of conditions are referred to in medicine as 'risk conditions' it is only causal risk factors that may constitute pathology. It also seems useful to consider a distinction between risk factors in the organism and risk factors in the organism's environment. For a risk factor to count as pathological it must reasonably concern some trait token of an organism.

Given that the relevant risk conditions to consider are causal risk factors that are part of the organism rather than the organism's environment and that we know what the risk is a risk of, it is however still not clear what a relevant risk condition consists in. For the purpose of understanding this, I think the DPE-theory's framework may be useful. The model of dispositions that the theory offers may serve as a useful theoretical model, especially in combination with the survival- and reproduction-efficiency measures. According to the

DPE-theory, pathology arises when a certain SDP (specific disposition profile) significantly reduces the further life expectancy of a survival-exemplary CDP (complete disposition profile). Hence, on the DPE-theory, survival-pathology is about SDPs' bearing on further life expectancy. For the purpose of examining the issue of risk conditions, we may ask whether there are different ways for an SDP to reduce the further life expectancy of a survival-exemplary CDP, such that these different ways also define a distinction between pathology and risk conditions.

Second, with clear definitions of the relevant risk conditions, it will be possible to see whether there is something conceptually different about risk conditions in comparison to (other) states of pathology. If there is, we may discuss whether this difference legitimates a clear separation of risk conditions from pathology, or whether risk conditions are best described as a subclass of pathological states.

Third, the question of whether risk conditions are different from pathology or not highlights the question mentioned in the introductory chapter of whether health and pathology are absolute or comparative concepts. For example, Giroux (2015, 180-183) argues that the extensive focus on risk factors in epidemiology indicates that health and disease are not binary, but comparative, notions:

> 'Risk-based diseases', such as the paradigmatic cases of hypertension and hypercholesterolemia (as used as an indicator of the atherosclerosis process), form a continuum with normal states and thus have an equivocal and unclear status, located somewhere between the normal and the pathological. The difficulties encountered in the definition of the normal level of blood pressure, i.e., in drawing the demarcation line between normotension and hypertension, are paradigmatic of this new equivocal status, which is in large part the result of modern epidemiology's risk approach. (2015, 181)

Fourth, the issue of how to conceive of risk conditions connects to other areas of discussion within the philosophy of medicine. One such area is that of overmedicalization. If risk conditions cannot be separated from pathology, one may worry that the requirements for health becomes too demanding. For example, Dawber, a prominent epidemiologist, writes that:

> Better knowledge of the natural history of the atherosclerotic process has led to a different concept of normality: that the normal person is one who not only has no disease but also is unlikely to develop it. At the extreme of this normality is the ideal individual who will *never* develop disease. The importance of this

> changing definition is best illustrated by the concept of risk factors as they pertain to the development of atherosclerotic disease. (1980, 223, italics are original)

Giroux (2015) discusses this quote and comments that "Dawber's approach raises some novel issues, for it seems to open the door to an unlimited broadening of the pathological domain and to the medicalisation of normal life" (2015, 182).

Another area of discussion to which the issue of how to conceive of risk conditions connects is that of prevention. Parallel to the terminological distinction between 'risk condition' and 'pathology' there is a terminological distinction between 'treatment' and 'prevention'. Whereas treatment is often considered an appropriate intervention in cases of pathology, prevention is often considered an appropriate intervention in cases of risk conditions. If it turns out that risk conditions cannot be separated from pathology, then it may also be questioned whether treatment and prevention are really different types of interventions.

7.5 Summary

In this chapter I have discussed four types of states, for which it is contested whether they should count as pathological. First, I discussed conditions that are trivial for the organism's survival or reproduction (e.g. a single malfunctioning cell). I defended the implication of the disposition profile efficiency theory that such conditions are not pathological.

Second, I distinguished between two types of normal ageing: on the one hand, functional impairments that inevitably occur among more or less all individuals while getting older (e.g. presbyopia), and, on the other hand, conditions that merely become more common in groups of individuals as they get older (e.g. senility of the Alzheimer type among people over 85). I defended the implication of the DPE-theory that the latter but not the former of these types of normal ageing are pathological.

Third, I defended the judgement that the activation of defense mechanisms (e.g. fever and vomiting) is typically not pathological, but healthy. I used the DPE-theory's model of dispositions to clarify why this is so.

Fourth, I discussed risk conditions (e.g. hypertension and osteoporosis). Here I did not argue for any specific position, but rather concluded that the DPE-theory might be useful for investigating into how risk conditions relate to pathology.

8 Conclusion

The aim of this thesis has been to develop a theory of health and pathology from a purely medical-theoretical point of view. The result is the disposition profile efficiency theory (abbreviated 'the DPE-theory'). The approach of this theory is explicative. It aims at definitions of 'health' and 'pathology' in physiology and pathology that are coherent, clear, and theoretically fruitful. The properties that the definitions of the DPE-theory refer to are natural properties. These properties need not, however, reflect natural kinds. As described in the introductory chapter, the definitions proposed are supposed to apply to somatic conditions, not only in humans, but in all organisms.

The DPE-theory distinguishes between health and pathology pertaining to survival, and health and pathology pertaining to reproduction. Roughly, it defines 'survival-health' of a trait token (e.g. an organ) as follows:

> A trait token is **survival-healthy** if and only if each of its specific disposition profiles (SDP) for its trait type's physiological functions has a high enough survival-efficiency.

However, this definition does not take into consideration that a specific disposition profile may have several different survival-efficiencies. The reason that it might have several different survival-efficiencies is that its survival-efficiency is relative (i) to a reference class to which the trait token's bearer belongs, and (ii) to a reference class-relative health standard (i.e. a survival-exemplary complete disposition profile). It is theoretically possible that an organism belongs to more than one reference class, and that there are more than one health standard for this reference class. Taking this into consideration, a more accurate definition of 'survival-health' of a trait token is the following:

> A trait token is **survival-healthy** if and only if each of its specific disposition profiles (SDP) for its trait type's physiological functions has a high enough survival-efficiency, irrespectively of which of the trait token's bearer's reference classes and irrespectively of which of these reference classes' standards for survival-health is being considered.

The DPE-theory defines 'survival-pathology' of a trait token roughly as follows:

> A trait token is **survival-pathological** if and only if it for at least one of its type's physiological functions has an SDP that does not have a high enough survival-efficiency.

Similar to the definition of 'survival-health', a more accurate definition of 'survival-pathology' must take into consideration that a specific disposition profile may have several different survival-efficiencies. A more accurate definition is:

> A trait token is **survival-pathological** if and only if it for at least one of its type's physiological functions has an SDP that does not have a high enough survival-efficiency, irrespectively of which of the trait token's bearer's reference classes, and irrespectively of which of these reference classes' standards for survival-health is being considered.

'Reproduction-health' and 'reproduction-pathology' are defined likewise.[34]

The unpacking of these definitions was done in Chapters 3-5. Specific disposition profiles (SDPs) and complete disposition profiles (CDPs) were introduced in Chapter 3. An SDP is a model of a trait token's disposition for performing a specific physiological function. It reports how (by what feature values for the relevant feature types) the trait token is disposed to perform a physiological function in different situations. A CDP represents the physiological functioning of a whole organism by modelling a collection of SDPs, one for each of the individual's trait tokens' type's physiological functions.

A standard for health was discussed in Chapter 4. Improving on Hausman's (2012) idea of a standard for health, I accounted for the standard for survival-health as a survival-exemplary CDP, relative to a reference class. A survival-exemplary CDP of a reference class is a CDP that maximizes further life expectancy, subject to certain constraints: the survival-exemplary CDP must be readily attainable for a significant share of the reference class members in common environments of these members. The standard for reproduction-health was accounted for likewise.

[34] All definitions presented in the previous chapters are collected in the appendix.

An account of reference classes was also presented in Chapter 4. The idea of the account is to search for a best overall fitting division of organisms into reference classes in accordance with five sorts of ideas. These ideas regard, first, structure and functionality; second, reference class-relativism; third, stability; fourth, age; and fifth, evolutionary aspects.

The accurate definitions of 'survival-health' and 'survival-pathology', and 'reproduction-health' and 'reproduction-pathology' above also included manipulations of survival-exemplary CDPs. Manipulations of survival- and reproduction-exemplary CDPs were presented in Chapter 5. A manipulation of a survival-exemplary CDP with an SDP dp consists in changing the survival-exemplary CDP so that its SDP corresponding to dp is set to dp. A manipulation of a reproduction-exemplary CDP works the same way.

Chapter 5 also discussed survival- and reproduction-efficiency. The basic idea is that the survival-efficiency of a trait token's SDP dp consists in a comparison between, on the one hand, the further life expectancy of a survival-exemplary CDP for a reference class of the trait token's bearer and, on the other hand, the further life expectancy of the same survival-exemplary CDP manipulated with dp. I proposed to carry out this comparison by division. So the survival-efficiency of an SDP dp is the ratio between the further life expectancy of the survival-exemplary CDP manipulated with dp and the further life expectancy of the survival-exemplary CDP. Likewise for reproduction-efficiency. An efficiency of 1 means that the SDP meets the standard. A value < 1 means that the SDP does not meet the standard. And a value > 1 means that the SDP is above the standard.

Chapter 5 further discussed how to understand "a high enough survival-efficiency", i.e. how to draw the line between survival-health and survival-pathology. I represented the distinguishing line by k_s, a constant real number. A survival-efficiency above k_s indicates health, whereas an efficiency below or equal to k_s indicates pathology. The corresponding line distinguishing reproduction-health from reproduction-pathology was represented by k_r. The values of k_s and k_r, I explained, are conventional, but within certain limits. They should be smaller than 1 and greater than 0. Further, they should reasonably be closer to 1 than to 0.

The DPE-theory was developed as an alternative to Boorse's biostatistical theory (the BST). The BST is the most debated theory of health and pathology, and its aim is also similar to that of the DPE-theory. I discussed the BST in depth in Chapter 2. There I pointed out several ways in which this theory is unclear. I also discussed some objections against the BST that I found im-

portant for the purpose of the thesis: arguments against the BST's individuation of reference classes; arguments against Boorse's choice of survival and reproduction as the relevant physiological goals; Kingma's dilemma; and the problem of common diseases.

In Chapter 1, I listed a number of desiderata that the sort of theory which this thesis aims at should fulfill. In order for the DPE-theory to count as successful, it should hence fulfill these desiderata. It should also fulfill them better than does the BST.

The first desideratum is that the theory should be theoretically sound. A reason for explicitly stating this as a desideratum is that theoretical soundness may come in conflict with making the theory practically applicable. The DPE-theory aims to be theoretically sound, but sometimes at the price of making the theory very difficult to apply in practice with precision. For example, we probably do not know about all feature types that matter for further life expectancy, or further expected number of offspring, for all trait types' physiological functions. Also, it will be very difficult and time consuming to find the individuations of reference classes in accordance with the account of reference classes presented in Chapter 4. Further, it will be difficult to calculate the exact survival-efficiencies of different dispositions, since it is hard to determine exact further life expectancies of different CDPs. This does, however, not mean that the DPE-theory is practically useless. The theory could be applied in a less precise sense. Even if we do not know everything about the relevant feature types for all different trait types' physiological functions, we know quite a lot. And even if it is very difficult to calculate exactly how organisms are divided into reference classes in accordance with the principles suggested in Chapter 4, we should be able to make such individuations somewhat roughly, using our observations of how different organisms function. Further, even if it is extremely difficult to determine exact further life expectancies, it should be possible for us to do qualified rough calculations of further life expectancies. Here we may make a parallel to utilitarianism. According to utilitarianism, an action is morally right if and only if it maximizes utility. That it is practically difficult for us to determine the utility of different possible actions does not mean that utilitarianism is not a useful theory. It offers an explanation of moral rightness, and although we cannot completely reliably determine the exact utility of different possible actions, we may make qualified rough estimations.

Also Boorse, I think, can be said to favor theoretical soundness over practical applicability. Is the DPE-theory more theoretically sound than the BST? The answer to this question is largely answered by the comparisons below of

how well the DPE-theory and the BST meet the remaining desiderata. We will see that the DPE-theory generally fares better in this respect.

The second desideratum is that the theory should not use any "non-natural" terms, but only empirical, statistical and logical terms. Like the BST, the DPE-theory does so. An interesting observation to make here, though, is that even if the DPE-theory is not statistical in the same sense as the BST, it still comprises statistical elements. The DPE-theory's standard for health was not purely statistically defined like the BST's, according to which the standard for health equals the average efficiency in the reference class. Yet, the DPE-theory's standard comprises statistical elements: the exemplary CDP is a CDP that is readily attainable for a *significant share* of the reference class members in *common environments* of the reference class members. Also, the account of reference classes contains similar statistical elements.

The third desideratum is that the theory should not itself involve values. It was explicitly argued that the DPE-theory is not value-involving with regard to its choice of goals, or its choice of reference classes. The question remains, however, whether the definitions that I have suggested are interesting and morally acceptable to use. This is a question which, I think, is difficult to answer in a definitive way. Rather, we should always be aware that the concepts we use within science may not be the most interesting, and that it may be morally dubious to use them. Whether they are interesting and morally fine to use may vary with the time, or scientific context, in which they are used. Hence the question of interest and ethical appropriateness of the definitions I have proposed is one to be revisited over and over again.

Like the DPE-theory, the BST does not itself involve any value. Yet, in comparison to the BST, I think that the DPE-theory is more realistic both in its motivation for survival and reproduction as the only relevant physiological goals and in its account of reference classes. This is, first, because the DPE-theory does not assume that the goals of survival and reproduction are objectively justifiable with reference to a phenomenon of intrinsic goal-directedness. It is, second, because in comparison to the BST the DPE-theory's account of reference classes offers more of an explanation why comparison classes of organisms tend to look, or should look, in a certain way in health evaluations.

The fourth desideratum is that the theory should be clear. The DPE-theory is clearer than the BST. In particular, it is clearer by its models of dispositions for performing physiological functions, in its account of efficiency, and in its account of the line-drawing between health and pathology. With regard to

these notions, it is also clearer than the theory of health and pathology suggested by Hausman (2012). However, there are parts of the DPE-theory that could be further clarified. In particular, in the account of survival- and reproduction-exemplary CDPs, it could be further elaborated how to define the set of possible worlds that determine whether a CDP is readily attainable for an individual.

The fifth desideratum is that the theory should (1) account for health and pathology as reference class-relative properties, and simultaneously (2) account for the importance of the integration of different physiological functions for health. Since Boorse leaves it a rather open question how to understand "efficiency", it is unclear whether we should understand the BST as accounting for health and pathology as reference class-relative properties or not. Because of the BST's vaguely defined notion of efficiency it is also unclear whether the BST accounts for the importance of the integration of different physiological functions for health, and if so, how.

The DPE-theory accounts for health and pathology as reference class-relative properties. It does so by the reference class-relative standards for health – survival- and reproduction-exemplary CDPs. Important in relation to this is also the manipulation, which lets us consider how a particular SDP affects the further life expectancy of a survival-exemplary CDP. The DPE-theory simultaneously accounts for the importance of the integration of different physiological functions for health. It does so by its models of dispositions. A trait token's SDP describes how the trait token performs its type's physiological function in different situations. These situations are not limited to the environment external to the organism. Rather, the situations contain the internal bodily environment of the trait token as well, i.e. effects of performances of other physiological functions. This means that an SDP describes how the trait token, in its disposition for performing a physiological function, is integrated with dispositions for other physiological functions in the organism.

The sixth desideratum is that the theory should have reasonable implications. Here I specifically brought up three types of conditions: common diseases (which should count as pathological), situation-specific normal functioning (which should count as healthy), and situation-specific diseases (which should count as pathological). We saw that the BST correctly classifies states of situation-specific normal functioning as healthy and states of situation-specific diseases as pathological. However, the BST does not classify states of common diseases as pathological.

In Chapter 6, I discussed what the DPE-theory says about these three types of conditions. I explained that on the DPE-theory, in contrast to the BST, states

of common diseases are pathological. This is because the DPE-theory (i) defines a standard for health distinct from the actual average in the reference class, and (ii) defines the line distinguishing health from pathology without reference to the distribution of efficiencies in the reference class.

I also explained that the DPE-theory, like the BST, classifies states of situation-specific normal functioning as healthy. This is because, on the DPE-theory, it is dispositions that are relevant for health, not single performances. The theory determines the health status of a trait token by comparing its dispositions to the corresponding ones of the exemplary CDP.

Further, I explained that the DPE-theory, like the BST, classifies states of situation-specific diseases as pathological. This is because, on the DPE-theory, health evaluations are not dependent on the particular situation of the trait token under evaluation. Whether a trait token is survival-healthy or survival-pathological only depends on how well its SDPs generally (i.e. not dependent on any particular situation) serve the goal of survival.

As we have now seen, the DPE-theory fulfills all of the desiderata well. It fulfills all of them at least as well as the BST, and several of them better. An issue that, however, requires attention is that of associated values. Although the DPE-theory is not by itself value-involving, we need to be cautious about the values guiding our interest in the relations described by the theory. Also, although the DPE-theory is much clearer than the BST, it should be acknowledged that the theory could be further clarified by investigations into how to account for the set of possible worlds that determine whether a CDP is readily attainable for an individual. Suggested future research also includes the relation between pathology and risk conditions. In Chapter 7, I discussed this issue, and suggested some more specific research questions. I hope to be able to deal with these questions in the near future.

Appendix: Definitions

Disposition Profiles

$RFT(F)$ is the set of relevant feature types for the physiological function F.

A **feature value configuration** for a physiological function F is a function

$fvc : RFT(F) \rightarrow$ values

such that for each feature type T in $RFT(F)$, $fvc(T)$ is a value for T.

A **specific disposition profile** for a physiological function F is a function

$dp :$ situations $\rightarrow$ feature value configurations

such that $dp(S)$ is a feature value configuration for F for each situation S.

A **complete disposition profile** for a set P of physiological functions is a function

$DP : P \rightarrow$ specific disposition profiles

such that $DP(F)$ is a specific disposition profile for the physiological function F for each F in P.

Reference Class-Relative Standards for Health

A CDP DP is **survival-exemplary**, relative to a reference class R, if and only if

out of the CDPs that are readily attainable for a significant share of the organisms in R in common environments for the organisms in R, DP gives the highest further life expectancy.

A CDP DP is **reproduction-exemplary**, relative to a reference class R, if and only if

out of the CDPs that are readily attainable for a significant share of the organisms in R in common environments for the organisms in R, DP gives the highest further expected number of offspring.

Reference Classes

An organism O **belongs to the reference class** R, at the time interval t, if and only if R is assigned to O-at-t in accordance with the division of all organisms that live at t or have lived not more than a certain number of generations back in time from t in accordance with Principles 1-6.

Principle 1. Each reference class must contain a sufficiently large number of individuals.

Principle 2a. For each reference class, there is exactly one set of organism part types with associated function types that together are readily attainable for a significant share of the reference class members in common environments of the reference class.

Principle 2b. Each of these functions contributes to the further life expectancy of a survival-exemplary CDP of the reference class, or to the further expected number of offspring of a reproduction-exemplary CDP of the reference class.

Principle 3. Each reference class only contains individuals of the same age.

Principle 4. No reference class includes individuals belonging to evolutionarily separated groups, i.e. individuals that do not have a close enough common ancestor.

Principle 5. Given Principles 1-4, the division of reference classes maximizes the total number of trait types among the individuals taken into consideration in the division.

Principle 6. The division of individuals generates no more reference classes than is needed to fulfill Principles 1-5.

Manipulations

$f(dp)$ is the physiological function that the SDP dp is an SDP for.

The **manipulation** of a CDP DP with an SDP dp is the CDP DP' which is exactly like DP except that $DP'\big(f(dp)\big) = dp$.

Efficiency

$\boldsymbol{R}(O,t)$ is the set of all reference classes to which the organism O belongs at the time interval t.

$\boldsymbol{S\text{-}Ex}(R)$ is the set of all survival-exemplary CDPs of the reference class R.

$\boldsymbol{R\text{-}Ex}(R)$ is the set of all reproduction-exemplary CDPs of the reference class R.

$fle(DP)$ is the further life expectancy of an organism with the CDP DP.

$feno(DP)$ is the further expected number of offspring of an organism with the CDP DP.

$M(DP, dp)$ is the manipulation of the CDP DP with the SDP dp.

Let O be an individual at the time interval t. Let dp be an SDP in the CDP representing O. Let $R \in \boldsymbol{R}(O,t)$ and let $DP_{S\text{-}Ex} \in \boldsymbol{S\text{-}Ex}(R)$. Provided that $fle(DP_{S\text{-}Ex}) > 0$, the **survival-efficiency** of dp relative to R and $DP_{S\text{-}Ex}$ is

$$s\text{-}eff(dp, R, DP_{S\text{-}Ex}) = \frac{fle\big(M(DP_{S\text{-}Ex}, dp)\big)}{fle(DP_{S\text{-}Ex})}$$

Let O be an individual at the time interval t. Let dp be an SDP in the CDP representing O. Let $R \in \boldsymbol{R}(O,t)$ and let $DP_{R\text{-}Ex} \in \boldsymbol{R\text{-}Ex}(R)$. Provided that $feno(DP_{R\text{-}Ex}) > 0$, the **reproduction-efficiency** of dp relative to R and $DP_{R\text{-}Ex}$ is

$$r\text{-}eff(dp, R, DP_{R\text{-}Ex}) = \frac{feno\big(M(DP_{R\text{-}Ex}, dp)\big)}{feno(DP_{R\text{-}Ex})}$$

Health and Pathology

A trait token a of an organism O is **survival-healthy** at a time interval t if and only if each of its SDPs dp for the physiological functions for a's trait type is such that, for each $R \in \boldsymbol{R}(O,t)$ and each $DP_{S\text{-}Ex} \in \boldsymbol{S\text{-}Ex}(R)$, $s\text{-}eff(dp, R, DP_{S\text{-}Ex}) > k_s$.

A trait token a of an organism O is **survival-pathological** at a time interval t if and only if at least one of its SDPs dp for a physiological function for a's trait type is such that, for each $R \in \boldsymbol{R}(O,t)$ and each $DP_{S\text{-}Ex} \in \boldsymbol{S\text{-}Ex}(R)$, $s\text{-}eff(dp, R, DP_{S\text{-}Ex}) \leq k_s$.

A trait token a of an organism O is **reproduction-healthy** at a time interval t if and only if each of its SDPs dp for the physiological functions for a's trait type is such that, for each $R \in \boldsymbol{R}(O,t)$ and each $DP_{R\text{-}Ex} \in \boldsymbol{R\text{-}Ex}(R)$, $r\text{-}eff(dp, R, DP_{R\text{-}Ex}) > k_r$.

A trait token a of an organism O is **reproduction-pathological** at a time interval t if and only if at least one of its SDPs dp for a physiological function for a's trait type is such that, for each $R \in \boldsymbol{R}(O,t)$ and each $DP_{R\text{-}Ex} \in \boldsymbol{R\text{-}Ex}(R)$, $r\text{-}eff(dp, R, DP_{R\text{-}Ex}) \leq k_r$.

Svensk sammanfattning

Vad innebär det att ett organ är friskt? Vad innebär det att ett organ är sjukt, eller patologiskt? En grov analys, som i princip alla håller med om, är att ett organ är friskt om det fungerar tillräckligt bra och att ett organ är sjukt eller patologiskt om det inte fungerar tillräckligt bra. Oenighet råder om hur man ska förstå ”fungerar” och hur man ska förstå ”tillräckligt bra”.

Funktioner, hälsa, och sjukdom har diskuterats sedan Aristoteles. Den moderna debatten om dessa begrepp tilltog i och med Wrights nydanande analys av funktionsbegreppet (1973,161, här översatt till svenska):

> Funktionen hos *X* är *Z betyder att*
> (*a*) *X* finns där eftersom den/det gör *Z*,
> (*b*) *Z* är en konsekvens (eller ett resultat) av att *X* finns där.

Denna idé om funktioner satte igång en intensiv debatt om funktioner i allmänhet, men också en debatt om biologiska, eller fysiologiska, funktioner i synnerhet. Detta stimulerade i sin tur debatten om hälsa och sjukdom från ett biologiskt perspektiv: Nya idéer om fysiologiska funktioner gav direkt idéer om hur vi bör förstå funktionsbegreppet i ”fungerar tillräckligt bra”, och bidrog också indirekt till frågan om hur ”tillräckligt bra” bör förstås. Om vi till exempel tittar på Boorses (1976b) mål-analys av fysiologiska funktioner, så bestäms fysiologiska funktioner av deras bidrag till överlevnad eller reproduktion. Ett hjärta har funktionen att pumpa blod eftersom blodpumpandet bidrar till organismens överlevnad. Enligt denna analys blir det naturligt att förstå ”tillräckligt bra” i termer av tillräckligt bidrag till överlevnad eller reproduktion. Ett hjärta är friskt om dess blodpumpande bidrar tillräckligt mycket till organismens överlevnad, och det är patologiskt om det bidrar otillräckligt mycket till organismens överlevnad.

Debatten om sjukdom som kom igång efter Wright (1973) kom till en början att mestadels kretsa kring motsättningen mellan ett biologiskt kriterium (om dysfunktion) och ett normativt kriterium (om en negativ värdering). Enligt Boorse (1977), en av pionjärerna på området, så kan ’sjukdom’ definieras som ett rent teoretiskt, värdefritt begrepp. Enligt många andra (t.ex. Agich

1983; DeVito 2000; Engelhardt 1975; Hare 1986; Reznek 1987; Wakefield 1992) är sjukdomsbegreppet oundvikligen värdeladdat. Även om debatten idag är mer nyanserad, så är motsättningen mellan en värdefri och en värdeladdad syn på hälsa och sjukdom fortfarande ett pågående tema i debatten (Kingma 2017).

Denna avhandling undersöker hälsa och patologi som teoretiska begrepp i biologi, eller mer specifikt medicinsk teori (d.v.s. fysiologi och patologi). Jag använder termen 'patologi' snarare än 'sjukdom', eftersom detta är en mer teoretiskt orienterad term. Liksom de flesta inom debatten så ämnar jag att med denna term täcka in alla tillstånd som inte är friska, d.v.s. infektioner såväl som brutna ben och medfödda missbildningar. Syftet med avhandlingen är att explikera 'hälsa' och 'patologi' som teoretiska termer inom medicinsk teori, d.v.s. att definiera dessa termer klart och tydligt på ett sätt som gör dem koherenta och fruktbara som teoretiska begrepp inom denna disciplin.

Definitionerna är tänkta att referera till naturliga egenskaper. Dessa egenskaper behöver emellertid inte reflektera naturliga sorter. Det bör också sägas att definitionerna siktar på somatiska tillstånd, inte bara hos människor, utan hos alla organismer. Eftersom det handlar om biologiska termer tänkta att referera till naturliga egenskaper, så är det rimligt att definitionerna är värdefria. De bör endast innehålla empiriska, statistiska, och logiska termer.

Även om det är önskvärt att definitionerna är praktiskt tillämpbara, så är det inte det primära målet. Det primära målet är teoretisk sundhet. Huvudskälet till att jag prioriterar teoretisk sundhet framför praktisk tillämpbarhet är att frågan jag är intresserad av är huruvida det är möjligt att definiera 'hälsa' och 'patologi' på ett sätt som är rimligt med avseende på medicinsk teori. Detta är en mer grundläggande fråga än huruvida termerna kan definieras på ett praktiskt tillämpbart sätt. För att kunna formulera rimliga praktiskt tillämpbara definitioner måste man först söka klarhet i huruvida dessa faktiskt refererar till det de är tänkta att referera till.

Resultatet av denna undersökning är vad jag kallar 'the disposition profile efficiency theory', eller (klumpigare) på svenska 'dispositionsprofileffektivitetsteorin'. Jag förkortar den 'DPE-teorin'. Fyra utmaningar för den typ av teori som DPE-teorin är, enligt beskrivningen ovan, är följande: (i) teorin bör vara klar och tydlig, (ii) teorin bör ha rimliga implikationer, (iii) teorin bör (som jag nämnt ovan) vara värdefri, och (iv) teorin bör samtidigt som den redogör för hälsa och patologi som referensklass-relativa egenskaper redogöra för vikten av integration av olika fysiologiska funktioner inom en individ för hälsa. Dessa fyra utmaningar tar jag upp eftersom de antingen formulerats i debatten med avseende på den mest omskrivna teorin om hälsa och patologi

som teoretiska begrepp i medicinsk teori – Boorses ”biostatistiska teori” – eller för att de, även om de inte diskuterats tidigare, är viktiga utmaningar som denna teori står inför. För att DPE-teorin ska kunna betraktas som framgångsrik behöver den klara samtliga fyra utmaningar bättre än Boorses biostatistiska teori.

I denna sammanfattning redogör jag kortfattat för avhandlingens mest centrala delar. Jag beskriver först Boorses biostatistiska teori. Sedan diskuterar jag de fyra utmaningarna och hur dessa drabbar denna teori. Jag presenterar sedan DPE-teorin i grova drag och diskuterar hur den klarar de fyra utmaningarna. Vi kommer att se att DPE-teorin klarar samtliga fyra utmaningar bättre än den biostatistiska teorin. Slutligen tillhandahåller jag, för den som vill läsa vidare i avhandlingen, en beskrivning av hur avhandlingen är disponerad i olika kapitel.

Boorses biostatistiska teori

Boorses biostatistiska teori (förkortad ’BST’) (se t.ex. Boorse 1977; 1997: 2014) är uppbyggd av fyra komponenter: (i) referensklasser, (ii) fysiologiska funktioner, (iii) effektivitet, och (iv) statistisk normalitet. Kortfattat definierar BST ’hälsa’ och ’patologi’ på följande sätt: Ett organ i en organism är friskt om och endast om organet är beskaffat att utföra sina *fysiologiska funktioner* vid typiska tillfällen med en *effektivitet* som är åtminstone *statistiskt normal* för den typen av organ i organismens *referensklass*. Annars är organet patologiskt. Referensklasser består enligt Boorse av individer av samma art, kön och ålder (t.ex. 7-9-åriga flickor). Fysiologiska funktioner är utföranden som hos individerna i en referensklass generellt sett bidrar till överlevnad eller reproduktion (såsom hjärtats blodpumpande). Effektivitet handlar om hur mycket ett utförande av en fysiologisk funktion bidrar till överlevnads- eller reproduktionschanser. Ett utförande med åtminstone statistiskt normal effektivitet är ett utförande vars effektivitet inte är alltför låg i förhållande till medeleffektiviteten för den typen av utförande inom referensklassen. ”Alltför låg i förhållande till medeleffektiviteten” avser här att utförandet, i den statistiska distributionen av utföranden i referensklasses, befinner sig långt under medeleffektiviteten. Detta innebär att en 8-årig flicka, vars hjärta är beskaffat att pumpa blod med en medeleffektivitet för 7-9-åriga flickor, har ett friskt hjärta. En annan 8-årig flicka, vars hjärta är beskaffat att pumpa blod med en effektivitet långt under denna medeleffektivitet, har däremot ett patologiskt hjärta.

Fyra utmaningar

Låt oss nu titta närmare på de fyra utmaningarna om klarhet, rimliga implikationer, värdefrihet, samt referensklass-relativism och integration, och se i vilken utsträckning BST klarar dessa.

Klarhet

Jag väljer att ta upp utmaningen om klarhet eftersom BST bitvis är relativt oklar. Framför allt är det oklart hur man ska förstå effektivitetsbegreppet som Boorse använder: I vilken enhet mäts överlevnads- och reproduktionschanser? Och givet en mätenhet för detta, hur mäts effektiviteten hos ett organs utförande av en fysiologisk funktion?

Rimliga implikationer

Utmaningen om rimliga implikationer har lyfts fram eftersom det har verkat som att teorier, sådana som BST, får orimliga implikationer gällande vissa typer av tillstånd. En sådan typ av tillstånd är vanliga sjukdomar. Schwartz (2007, 375) ger några exempel på vanliga sjukdomar som tveklöst klassificeras som patologiska inom medicinsk teori. En sådan sjukdom är höftledsdysplasi, en dysfunktion hos höftleden vanlig bland hundar. För vissa raser förekommer denna dysfunktion hos 30 procent av individerna. Om vi tittar på människor, så är ett annat exempel godartad prostataförstoring, ett tillstånd som försämrar förmågan att urinera. Detta tillstånd uppskattas förekomma hos 17 procent av alla män över 70 år. Ytterligare ett exempel är Alzheimers demens som drabbar 16 procent av alla människor äldre än 85 år. Anledningen till att debatten har fokuserat på vanliga sjukdomar är att BST inte klarar att redogöra för denna typ av tillstånd som patologiska. Detta beror på att Boorse skiljer hälsa från patologi rent statistiskt. Den statistiska gränsdragningen mellan hälsa och patologi innebär att specifika typer av patologiska tillstånd per definition måste vara ovanliga. Schwartz exempel på vanliga sjukdomar visar att denna implikation av BST inte är acceptabel. Snarare måste en sund teori om hälsa och patologi kunna klassificera även vanliga tillstånd som patologiska.

Två andra typer av tillstånd som pekats ut som problematiska är (i) vad som betraktas som friska situationsspecifika utföranden av fysiologiska funktioner, och (ii) situationsspecifika sjukdomar. Kingma (2010) har poängterat att vissa fysiologiska funktioner är adekvata att utföra enbart i vissa situationer. Till exempel är det adekvat för matsmältningssystemet att smälta mat när det finns mat i det, men inte annars. Och för många organ eller fysiologiska system så varierar den adekvata intensiteten av utförandet med situationen. Även om det

alltid är adekvat för hjärtat att pumpa blod så skiljer sig den adekvata pulsen åt mellan olika situationer. Den skiljer sig exempelvis mellan situationer där individen vilar, går, eller springer. Detta innebär att en teori om hälsa och patologi måste kunna klassificera ett organ som friskt även om det inte utför en fysiologisk funktion maximalt, givet att det befinner sig i en situation där det inte är adekvat att utföra funktionen i fråga maximalt.

Det gäller också för många tillstånd som klassificeras som patologiska inom medicinsk teori att de är typiska i specifika situationer. Kingma (2010) ger följande exempel: leversvikt vid överdosering av paracetamol, skörbjugg vid vitamin C-brist, och frakturer vid yttre våld. Som Kingma poängterar så utgör dessa tillstånd paradigmatiska fall av patologi. En teori om hälsa och sjukdom som begrepp i medicinsk teori bör därför kunna klassificera situationsspecifika sjukdomar som patologiska. Kingma (2010) argumenterar dock för att BST, och teorier av samma typ, inte kan redogöra både för friska situationsspecifika utföranden av fysiologiska funktioner och för situationsspecifika sjukdomar. Kingma riktar alltså primärt sin kritik mot BST, som enligt beskrivningen ovan definierar ’hälsa’ och ’patologi’ i termer av statistisk normal funktion. I grova termer ser hennes invändning ut som följer. För att BST ska kunna redogöra för friska situationsspecifika utföranden av fysiologiska funktioner, så måste vad som räknas som ett friskt utförande betraktas relativt situation. Så ett friskt utförande i en viss situation bestäms av vad som är ett typiskt utförande *i den situationen*. Detta, argumenterar Kingma, innebär att det inte blir möjligt för BST att redogöra för patologiska tillstånd som är typiska i specifika situationer. Även om tillståndet i fråga är otypiskt generellt sett, så är det typiskt givet situationen – och måste alltså enligt BST räknas som friskt.

Kingmas invändning håller dock inte. Som poängterats, framför allt av Hausman (2011), så bygger den på en missuppfattning av BST. Den springande punkten är att Kingma inte skiljer på specifika situationers normala påverkan på organ å ena sidan och organs normala beskaffenheter, eller dispositioner, att utföra fysiologiska funktioner relativt situationer å andra sidan. Till exempel så är leversvikt en normal påverkan på levern vid överdosering av paracetamol, men de nya dispositioner som levern får i och med leversvikten är inte typiska att ha. Vad som är relevant för hälsa och patologi är just det senare, d.v.s. organs dispositioner att utföra fysiologiska funktioner, och inte huruvida orsaken till att dessa uppkommit skulle generera samma dispositioner hos levern hos andra individer i samma referensklass. Så länge överdosering av paracetamol inte är typiskt i en referensklass så är det inte typiskt – och därmed inte heller friskt – att ha leversvikt.

Även om BST undkommer Kingmas invändning, så har teorin problem att redogöra för all situationsspecificitet hos fysiologiska funktioner. Eftersom hälsa och patologi enligt Boorses definition enbart handlar om utförande av fysiologiska funktioner *vid typiska tillfällen*, så kan BST inte redogöra för sjukdomar som består i att ett organ eller fysiologiskt system utför en fysiologisk funktion även i situationer där det adekvata varit att inte utföra funktionen över huvud taget. Exempel på detta är autoimmuna sjukdomar där immunförsvaret reagerar även mot kroppsegna celler, d.v.s. även i situationer där det inte borde aktiveras.

Värdefrihet

Även om Boorse hävdar att BST är värdefri, så har teorin har anklagats för att vara värdeinvolverande. BST innehåller inga värdeladdade termer, och har inte heller några normativa implikationer i sig själv. Kingma (2007; 2014) har argumenterat för att teorin trots detta inte är värdefri. Kingma poängterar att, beroende på hur referensklasser individueras, så kommer olika tillstånd att räknas som friska eller patologiska. Till exempel så kommer extensionen av 'patologi' vara olika om referensklasser individueras enbart av art, kön och ålder jämfört med om referensklasser dessutom individueras av sexuell läggning, kromosomala varianter eller alkoholkonsumtion. Om referensklasser individueras enbart av art, kön och ålder så kommer homosexualitet, Downs syndrom och skrumplever räknas som patologiska tillstånd eftersom de reducerar chanserna för överlevnad eller reproduktion i för hållande till referensklassen. Om sexuell läggning, kromosomala varianter och alkoholkonsumtion däremot tas med i individueringen så kommer dessa tillstånd istället (ofta) att räknas som friska, eftersom de då blir normala i de referensklasser där de (oftast) förekommer. Kingmas poäng är sedan att det inte finns någon objektiv grund för att välja en viss individuering av referensklasser snarare en någon annan. Boorses val av referensklasser blir därmed värdeinvolverande, menar hon: det finns flera olika sätt att individuera referensklasser på och valet att individuera dem just genom art, kön och ålder är baserat på värdeomdömen svarare än objektiva fakta.

Även om Kingma har en poäng med sitt argument så stämmer det inte att Boorses val av referensklasser gör BST värdeladdad. Även om det är värdeomdömen som styr valet av referensklasser, så är själva teorin fortfarande värdefri. Värderingen styr vad i naturen vi väljer att studera med begreppen hälsa och patologi, men det vi studerar blir inte därmed värdeladdat. Att det ligger en värdering bakom valet av studieobjekt är gemensamt för alla vetenskapliga begrepp. Fysiker, till exempel, beskriver olika partiklar och deras relationer

eftersom sådan kunskap värdesätts. Någon skulle kanske vilja invända här att protoner, elektroner och andra partiklar, till skillnad från hälsa och patologi, är naturliga sorter. Detta spelar emellertid ingen roll. Även om fysiker intresserar sig för olika partiklar för att de är naturliga sorter, så är det ändå en värdering som styr vad de forskar om – en värdering knuten till naturliga sorter. Om det som BST beskriver med sina definitioner av ’hälsa’ och ’patologi’ är intressant, så är det så på grund av vissa värderingar. Dessa värderingar är dock inte essentiellt annorlunda än de värderingar som motiverar fysikers intresse för olika partiklar.

Även om BST inte kan anklagas för att vara värdeladdad på grund av att teorin beskriver vissa saker i naturen snarare än andra så kan man ändå ifrågasätta huruvida fysiologer och patologer bör intressera sig för det som BST’s definitioner beskriver. Man kan vara skeptisk till de värderingar som motiverar beskrivningar av hälsa och patologi, så som Boorse redogör för dessa. Om valet av referensklasser inte reflekterar något fysiologiskt intressant, utan kanske till och med reflekterar moraliskt tvivelaktiga idéer, t.ex. om sexuell läggning, kan det vara på sin plats att argumentera för en ändring av hur referensklasser individueras. Boorse motiverar indelningen av referensklasser genom art, kön och ålder framför allt med påståendet att det verkar vara så referensklasser är indelade i fysiologi. Men han ger ingen djupare motivering till detta sätt att individuera referensklasser.

Referensklass-relativism och integration

Den fjärde utmaningen handlade dels om referensklass-relativism, dels om integration av fysiologiska funktioner. Jag argumenterar i avhandlingen för att hälsa och patologi bör redogöras för som referensklass-relativa egenskaper. Detta innebär att huruvida ett specifikt organ är friskt eller patologiskt bestäms av en jämförelse av organet med en standard som är gemensam för alla individer i organismens referensklass. Av detta följer att om två hjärtan hos två olika individer tillhörande samma referensklass är exakt likadana (d.v.s. disponerade att utföra hjärtats fysiologiska funktioner på precis samma sätt), så har de samma hälsostatus. Detta gäller även om hjärtat, så som det är disponerat, fungerar bättre i den ena individen än den andra, på grund av att de två individerna i övrigt är olika disponerade. Varför bör hälsa och patologi redogöras för på detta sätt? Till att börja med så finner vi indikationer på detta i medicinsk teori. Inom fysiologi beskrivs friskt fungerande organ för grupper av individer, snarare än för enskilda individer. Och inom patologi beskrivs mer eller mindre typiska nedsättningar, resulterande i patologi, hos dessa typer

av organ. Även om det finns meningsskiljaktigheter om huruvida hälsostatusen hos ett organ endast är relativ organismen i fråga eller om den är relativ organismens referensklass, så finns det åtminstone vissa intuitioner som lutar åt att ett organs hälsostatus är densamma oavsett vilken individ (inom en referensklass) organet tillhör. Och även om termerna ’hälsa’ och ’patologi’ ibland används på ett individ-relativt sätt, så är det meningsfullt att skilja hälsa från patologi på ett referensklass-relativt sätt. Säg till exempel att en individ har en reducerad kapacitet hos organet $o1$, men att denna reduktion kompenseras helt av en atypisk hög kapacitet hos organet $o2$. Enligt en individ-relativ syn på hälsa och patologi är $o1$ friskt, och enligt en referensklass-relativ syn är $o1$ istället patologiskt. Poängen är nu att även om vi antar ett individ-relativt synsätt och bedömer $o1$ vara friskt, så skulle det ändå vara intressant att veta att en anledning till att $o1$ inte är patologiskt är att ett annat organ kompenserar (utöver ”vad som kan förväntas”) för den reducerade kapaciteten hos $o1$. Om hälsa och patologi förstås som en relation mellan ett mål, en individ och ett organ hos denna individ (vilket går att prata om relativt enkelt utan att använda termerna ’hälsa’ och ’patologi’) så skulle det inte finnas några (existerande) termer kvar för att prata om den mer komplexa relationen mellan ett mål, en individ, ett organ hos denna individ och denna individs referensklass – en relation som enligt indikationerna ovan verkar uppta åtminstone visst intresse i medicinsk teori.

Det kan tyckas som att en referensklass-relativ syn på hälsa får problem att ta i beaktande att integration av olika fysiologiska funktioner inom en individ är viktigt för hälsa. På en detaljerad nivå är i princip alla organismer unika, med utföranden av olika fysiologiska funktioner avvägda mot varandra. Som ett exempel kan nämnas människans system för reglering av energikonsumtion, vilket involverar sköldkörteln, hypofysen och hypotalamus. Regleringen sker genom hormonerna T3 och T4, vilka påverkar energiproduktionen i kroppens celler. Dessa hormoner utsöndras av sköldkörteln. Sköldkörtelns utsöndring av T3 och T4 stimuleras i sin tur av hormonet TSH, vilket utsöndras av hypofysen. Hypofysens utsöndring av TSH stimuleras åter i sin tur av hormonet TRH, vilket utsöndras av hypotalamus. Systemet innehåller också en negativ återkopplingsmekanism. Både hypotalamus och hypofysen är känsliga för cirkulerande T3 och T4 på så sätt att ju högre nivåer av T3 och T4 som cirkulerar i blodet, desto mindre TRH och TSH utsöndrar de. Säg att vi vill undersöka hälsostatusen hos en specifik hypofys. Det räcker då inte att bara titta på mängden TSH som den utsöndrar. Snarare måste vi titta på mängden TSH som den utsöndrar i relation till mängden TRH utsöndrat av hypotalamus samt mängden T3 och T4 utsöndrat av sköldkörteln.

En teori om hälsa och patologi som teoretiska begrepp inom medicinsk teori bör alltså både utvärdera hälsostatusen hos ett organ gentemot en referensklass-gemensam standard och samtidigt ta i beaktande att en viktig aspekt av ett organs hälsa är att dess utföranden av fysiologiska funktioner är väl integrerat med utföranden av andra funktioner inom individen. Redogör BST för hälsa och patologi som referensklass-relativa egenskaper? Eftersom Boorse, som jag påpekat ovan, lämnar det relativt öppet hur effektivitetsbegreppet som han använder bör förstås, så är det också oklart huruvida BST kan sägas redogöra för hälsa och patologi som referensklass-relativa egenskaper eller ej. I och med det endast vagt definierade effektivitetsbegreppet blir det också oklart huruvida, och i så fall hur, BST tar vikten av integration av fysiologiska funktioner inom en individ för hälsa i beaktande.

Om vi nu summerar kan vi konstatera att BST inte klarar samtliga fyra utmaningar: Till att börja med är teorin bitvis oklar, framför allt med avseende på effektivitetsbegreppet. För det andra ger teorin problematiska implikationer, dels när det handlar om vanliga sjukdomar och dels när det handlar om sjukdomar som består i att ett organ utför en fysiologisk funktion även i situationer då det adekvata vore att inte utföra funktionen alls. Den tredje utmaningen, om värdefrihet, klarar BST i princip, eftersom teorin inte i sig är värdeladdad. Men valet att studera det som BST's definitioner av 'hälsa' och 'patologi' pekar ut skulle kunna försvaras bättre om Boorse kunde motivera valet av referensklasser på ett mer grundläggande plan. Gällande den fjärde utmaningen, om referensklass-relativitet och integration, så är BST åter oklar. BST kan inte sägas utesluta att hälsa och patologi är referensklass-relativa egenskaper och att integration av fysiologiska funktioner inom en individ är väsentligt för hälsa. Men detta är inget som Boorse tar upp, och vi ges därmed inte heller något förslag på hur teorin skulle klara denna utmaning.

DPE-teorin

DPE-teorin, som jag utvecklar i avhandlingen, liknar BST på så sätt att den definierar 'hälsa' och 'patologi' med hänvisning till referensklasser, fysiologiska funktioner, effektivitet och en linjedragning mellan tillräckligt höga och för låga effektivitetsnivåer för hälsa. Men teorin skiljer sig också från BST på flera väsentliga punkter för att klara utmaningarna om klarhet, rimliga implikationer, värdefrihet, och referensklass-relativism och integration. Nedan kommer jag beskriva DPE-teorin i grova drag och förklara hur den hanterar de fyra utmaningarna.

Jag argumenterar i avhandlingen, liksom Boorse, för att överlevnad och reproduktion utgör de relevanta målen för fysiologiska funktioner. DPE-teorin

skiljer dock mellan hälsa och patologi med avseende på överlevnad, och hälsa och patologi med avseende på reproduktion. I denna sammanfattning kommer jag för enkelhetens skull att endast diskutera hälsa och patologi med avseende på överlevnad.

En grundbult i DPE-teorin är dispositionsprofiler. Till skillnad från Boorse, som tillskriver effektivitet till utföranden av fysiologiska funktioner, så tillskriver DPE-teorin effektivitet till, vad jag kallar, specifika dispositionsprofiler. Varje organ har, för var och en av dess typs fysiologiska funktioner, en specifik dispositionsprofil som beskriver hur organet är benäget att utföra (alternativt inte utföra) funktionen som respons på olika situationer. Jag kallar det en ”specifik” dispositionsprofil eftersom den endast hanterar en specifik fysiologisk funktion. Idén illustreras av nedanstående tabell som rapporterar om ett specifikt hjärtas benägenhet att pumpa blod (baserat på antagandet att utförande av denna funktion kan beskrivas fullständigt genom angivelse av slagfrekvens och kontraktionskraft):

	S_1	S_2	…	S_n	…
Slagfrekvens	v_1	v_2	…	v_n	…
Kontraktionskraft	v_1'	v_2'	…	v_n'	…

Här är varje S en situation, och varje v och v' värden som rapporterar om slagfrekvens respektive kontraktionskraft.

”Situation” här avser ett möjligt tillstånd av världen. Situationen inkluderar alltså inte bara sådant som ligger utanför den individ vars organ vi undersöker. Den inkluderar även allt inom individen. Situationerna i tabellen ovan innefattar alltså till exempel sådant som blodvolym, nervsignaler, samt olika hormonnivåer. På detta vis redogör DPE-teorin för vikten av integration av olika fysiologiska funktioner inom en individ för hälsa. Tänk igen på exemplet om människans system för reglering av energikonsumtion. Jag sa att för att kunna utvärdera hälsostatusen hos en specifik hypofys så räcker det inte att bara titta på mängden TSH som den utsöndrar. Snarare måste vi titta på mängden TSH som den utsöndrar i relation till mängden TRH utsöndrat av hypotalamus samt mängden T3 och T4 utsöndrat av sköldkörteln. Detta är inget problem om vi använder specifika dispositionsprofiler. Hos hypofysens specifika dispositionsprofil för utsöndring av TSH anges mängden TSH som hypofysen utsöndrar som svar på olika situationer, vilka innefattar specifika nivåer av TRH, T3 och T4.

Enligt DPE-teorin är ett organ friskt om och endast om samtliga av dess specifika dispositionsprofiler för dess typs fysiologiska funktioner har en tillräckligt hög effektivitet. Om och endast om någon av de specifika dispositionsprofilerna har en för låg effektivitet är organet patologiskt. Vad innebär det då att en specifik dispositionsprofil har tillräckligt hög, alternativt för låg, effektivitet? Ett svar på detta kräver till att börja med ett klargörande av hur effektivitet mäts.

För att mäta effektivitet behöver vi titta på, vad jag kallar, kompletta dispositionsprofiler. En komplett dispositionsprofil är en samling specifika dispositionsprofiler som, istället för att hantera endast en fysiologisk funktion, hanterar samtliga fysiologiska funktioner hos en individ. En komplett dispositionsprofil beskriver därmed hur en hel individ fungerar. För att illustrera idén, betänk en fiktiv organism som endast har ett hjärta och en njure. Låt oss anta att hjärtat endast har en fysiologisk funktion – att pumpa blod – och att njuren endast har två fysiologiska funktioner – att filtrera blod och att utsöndra hormonet renin (som påverkar blodtryck). Den kompletta dispositionsprofilen som representerar hur denna organism fungerar inkluderar en specifik dispositionsprofil för hjärtats blodpumpande, en specifik dispositionsprofil för njurens blodfiltrering och en specifik dispositionsprofil för njurens utsöndring av renin. Idén illustreras av nedanstående tabell (där varje dp är en specifik dispositionsprofil).

Pumpa blod	dp_1
Filtrera blod	dp_2
Utsöndra renin	dp_3

Vi kan jämföra hur bra två alternativa specifika dispositionsprofiler för samma fysiologiska funktion, dp och dp', är för en viss individs överlevnad genom att se hur individens överlevnadschanser påverkas när dp byts ut mot dp' i individens kompletta dispositionsprofil. Ett rimligt sätt att mäta överlevnadschanser är genom ytterligare förväntad livslängd. Vi kan därmed jämföra hur bra dp och dp' är för en viss individs överlevnad genom att undersöka hur individens ytterligare förväntade livslängd förändras när dp byts ut mot dp' i individens kompletta dispositionsprofil. Denna typ av jämförelser är, som vi kommer se nedan, väsentliga för att kunna bestämma effektiviteten hos specifika dispositionsprofiler.

Idén om effektivitet som jag presenterar i avhandlingen hänger ihop med tanken om referensklass-relativism. Som jag tog upp ovan bör en teori om hälsa och patologi som teoretiska begrepp inom medicinsk teori redogöra för

hälsa och patologi som referensklass-relativa egenskaper. Detta innebär för DPE-teorin att, oavsett vilket specifikt organ i vilken specifik individ tillhörande en viss referensklass som undersöks, så måste effektiviteten hos organets specifika dispositionsprofiler utvärderas i kontexten av en och samma kompletta dispositionsprofil. Den måste utvärderas i en komplett dispositionsprofil som utgör standarden för hälsa i referensklassen. Om en individ fungerar helt i enlighet med denna kompletta dispositionsprofil så är individen frisk i alla avseenden. Om en individ har ett organ som med avseende på någon specifik dispositionsprofil är betydligt sämre än motsvarande specifika dispositionsprofil hos denna kompletta dispositionsprofil så är organet patologiskt. ”Sämre” betyder här att organets specifika dispositionsprofil ger den kompletta dispositionsprofilen signifikant lägre ytterligare förväntad livslängd. Jag kallar den kompletta dispositionsprofil som utgör standarden för hälsa i en referensklass ’exemplarisk’.

Baserat på idén ovan, om hur vi kan jämföra två olika specifika dispositionsprofiler för en viss fysiologisk funktion, bestäms effektiviteten hos en specifik dispositionsprofil *dp* genom att se hur *dp* fungerar som alternativ specifik dispositionsprofil till den motsvarande specifika dispositionsprofilen i den exemplariska kompletta dispositionsprofilen. Följande exempel illustrerar idén. Vi är intresserade av att utvärdera effektiviteten hos människan Georgs röda benmärgs specifika dispositionsprofil för produktion av blodceller. Georg är 67 år gammal och lider av akut myeloisk leukemi. Detta innebär att hans röda benmärgs förmåga att producera vissa blodceller är gravt reducerad i förhållande till den exemplariska kompletta dispositionsprofilens för Georgs referensklass. Akut myeloisk leukemi är en allvarlig sjukdom som bland annat ger trötthet, feber, nattsvettningar, och förlängd koagulationstid. Utan behandling leder sjukdomen till döden inom loppet av några veckor eller månader. För att utvärdera effektiviteten hos Georgs röda benmärgs specifika dispositionsprofil för produktion av blodceller jämför vi den ytterligare förväntade livslängden för den exemplariska kompletta dispositionsprofilen för Georgs referensklass med den ytterligare förväntade livslängden för samma kompletta dispositionsprofil ändrad så att den istället för sin egen specifika dispositionsprofil för att producera blodceller har Georgs röda benmärgs specifika dispositionsprofil för denna funktion. Jag återkommer till detta exempel nedan.

Hur den exemplariska kompletta dispositionsprofilen för en referensklass ser ut beror på individerna som referensklassen utgörs av. Den exemplariska kompletta dispositionsprofilen ska vara relativt tillgänglig för en signifikant andel av referensklassens individer. Detta innebär att det måste gälla för en signifikant andel av individerna i referensklassen att individen antingen har

den kompletta dispositionsprofilen i fråga, eller att det finns minst en möjlig värld relativt nära den faktiska världen där individen har den kompletta dispositionsprofilen i fråga. Den kompletta dispositionsprofil som, utav de kompletta dispositionsprofiler som är relativt tillgängliga för en signifikant andel av individerna i referensklassen, ger den högsta ytterligare förväntade livslängden utgör den exemplariska kompletta dispositionsprofilen. I princip kan det finnas flera exemplariska kompletta dispositionsprofilen för en referensklass, men låt oss i denna sammanfattning för enkelhetens skull anta att det för varje referensklass finns exakt en exemplarisk komplett dispositionsprofil.

Hur bestäms då referensklasser? Redogörelsen för referensklasser som jag presenterar i avhandlingen är holistisk. Grundtanken är att söka efter den mest lämpliga indelningen av alla organismer, baserat på fem typer av idéer. Dessa idéer handlar om (i) struktur och funktionalitet, (ii) referensklass-relativism, (iii) stabilitet, (iv) ålder, och (v) evolutionära aspekter. De motiveras, dels, av grundläggande intressen inom fysiologi, dels, av teoretiska krav som följer med rollen som referensklasser spelar i DPE-teorin. Redogörelsen för referensklasser är därmed mer grundläggande än Boorses redogörelse, eftersom den ger mer av en förklaring till varför man typiskt tittar på, eller bör titta på, vissa referensklasser i hälsovärderingar. I och med detta lyckas DPE-teorin bättre än BST med att motivera varför det som ’hälsa’ och ’patologi’ refererar till enligt teorins definitioner är intressant och moraliskt godtagbart att studera.

En viss individ kan i teorin, enligt DPE-teorins redogörelse för referensklasser, tillhöra flera referensklasser. Låt oss i denna sammanfattning för enkelhetens skull anta att varje individ endast tillhör en referensklass.

För att sammanfatta teorin så långt, så beror ett organs hälsostatus på effektiviteten hos dess specifika dispositionsprofiler. Effektiviteten hos en specifik dispositionsprofil dp ges genom en jämförelse av dp med motsvarande specifika dispositionsprofil hos den exemplariska kompletta dispositionsprofilen för individens referensklass: vi undersöker hur den exemplariska kompletta dispositionsprofilens ytterligare förväntade livslängd påverkas om vi byter ut dess motsvarande specifika dispositionsprofil mot dp.

Jag föreslår i avhandlingen att vi utför denna jämförelse genom division. Säg att vi utvärderar effektiviteten hos en specifik dispositionsprofil dp. Vi har då å ena sidan den ytterligare förväntade livslängden för den exemplariska kompletta dispositionsprofilen för referensklassen som dp’s bärare tillhör. Kalla denna ytterligare förväntade livslängd $yfl(Ex)$. Vi har å andra sidan den ytterligare förväntade livslängden samma kompletta dispositionsprofil, men ändrad till att ha dp istället för dess egna motsvarande specifika dispositions-

profil. Kalla denna ytterligare förväntade livslängd $yfl(Ex, dp)$. Vi får effektiviteten för dp genom att dividera den andra ytterligare förväntade livslängden med den första. dp's effektivitet är alltså:

$$\frac{yfl(Ex, dp)}{yfl(Ex)}$$

Resultatet av denna division är ett icke-negativt reellt tal. Om värdet blir 1 innebär det att dp är lika bra för ytterligare förväntad livslängd som den motsvarande specifika dispositionsprofilen hos standarden. Ett värde < 1 innebär att dp är sämre för ytterligare förväntad livslängd än den motsvarande specifika dispositionsprofilen hos standarden. Ett värde > 1 innebär att dp är bättre för ytterligare förväntad livslängd än den motsvarande specifika dispositionsprofilen hos standarden.

Om vi nu igen tänker på exemplet med Georg som har akut myeloisk leukemi så bestäms effektiviteten för hans röda benmärgs specifika dispositionsprofil för produktion av blodceller som följer. Först behöver vi veta den ytterligare förväntade livslängden för den exemplariska kompletta dispositionsprofilen för Georgs referensklass. Låt oss anta att den är 20 år. Sedan behöver vi veta den ytterligare förväntade livslängden för samma kompletta dispositionsprofil men ändrad till att ha Georgs röda benmärgs specifika dispositionsprofil för produktion av blodceller istället för dess egna motsvarande specifika dispositionsprofil. Eftersom Georg har akut myeloisk leukemi skiljer sig hans röda benmärgs specifika dispositionsprofil för produktion av blodceller väsentligt från motsvarande specifika dispositionsprofil hos den exemplariska kompletta dispositionsprofilen. Låt oss anta att den ytterligare förväntade livslängden för den exemplariska kompletta dispositionsprofilen ändrad till att ha Georgs röda benmärgs specifika dispositionsprofil är 0,2 år, eftersom denna kompletta dispositionsprofil representerar en individ med akut myeloisk leukemi. Effektiviteten för Georgs röda benmärgs specifika dispositionsprofil för produktion av blodceller blir då $\frac{0{,}2}{20} = 0{,}01$. Georgs röda benmärgs specifika dispositionsprofil för produktion av blodceller har alltså effektiviteten 0,01.

Frågan som nu återstår är hur hög, respektive låg, effektivitet som krävs för hälsa och patologi. Jag representerar skiljelinjen mellan tillräckligt hög effektivitet och för låg effektivitet med en konstant. Effektiviteter som tar denna konstants värde, alternativt ett lägre värde, indikerar patologi medan effektiviteter som tar ett högre värde än denna konstant indikerar hälsa. Låt k vara en konstant (ett reellt tal) som representerar skiljelinjen mellan tillräckligt hög och för låg effektivitet. 'Hälsa' och 'patologi' definieras då på följande vis:

Ett organ a är **friskt** om och endast om varje specifik dispositionsprofil dp för a's organtyps fysiologiska funktioner har en effektivitet lägre än k.

Ett organ a är **patologiskt** om och endast om minst en av dess specifika dispositionsprofiler dp för någon av a's organtyps fysiologiska funktioner har en effektivitet lägre än eller lika med k.

Värdet för k är konventionellt, givet vissa ramar. Till att börja med är värdet lägre än 1. Annars skulle organ med specifika dispositionsprofiler utan minsta effektivitetsreduktion vara patologiska. Värdet är också högre än 0. Rimligen ligger värdet närmare 1 än 0. Annars skulle det krävas alltför omfattande nedsättning för att ett organ ska räknas som patologiskt. Om vi nu återgår till exemplet med Georg, vars röda benmärgs specifika dispositionsprofils effektivitet var 0,01, så kan vi konstatera att Georgs röda benmärg är patologisk. Värdet för k är rimligen högre än 0,01.

Jag har nu beskrivit DPE-teorin i grova drag. Vi har sett att teorin i enlighet med utmaningen om referensklass-relativism och integration både redogör för hälsa och patologi som referensklass-relativa egenskaper och redogör för vikten av integration av olika fysiologiska funktioner inom en individ för hälsa: Hälsa och sjukdom blir referensklass-relativa egenskaper i och med den referensklass-relativa standarden för hälsa. Den roll som integration av fysiologiska funktioner spelar för hälsa redogörs för genom de specifika dispositionsprofilerna, där utföranden av fysiologiska funktioner är relativa situationer vilka innefattar alla förhållanden inom individen.

Vi har också sett att DPE-teorins redogörelse för referensklasser motiveras på ett transparant sätt, med hänvisning till grundläggande intressen inom fysiologi samt teoretiska krav som följer med den roll referensklasser spelar i DPE-teorin.

Hur klarar DPE-teorin de övriga två utmaningarna, om klarhet respektive rimliga implikationer? DPE-teorin är relativt klar, åtminstone är den klarare än BST. Detta gäller framför allt med avseende på modellerna av dispositioner för utförande av fysiologiska funktioner och definitionen av 'effektivitet', men också med avseende på gränsdragningen mellan tillräckligt höga och för låga effektivitetsnivåer, samt redogörelsen för referensklasser.

Utmaningen om rimliga implikationer fokuserade på tre typer av tillstånd: (i) vanliga sjukdomar, (ii) friska situationsspecifika utföranden av fysiologiska funktioner, samt (iii) situationsspecifika sjukdomar. DPE-teorin hanterar vanliga sjukdomar dels genom hur standarden för hälsa definieras, dels

genom hur gränsen dras mellan tillräckligt hög och för låg effektivitet. Eftersom standarden inte bestäms rent statistiskt, d.v.s. som ett faktiskt medelvärde för referensklassen, utan i termer av vad som är *relativt tillgängligt* för en signifikant andel av individerna i referensklassen, kan standarden ligga högre än det faktiska statistiska medelvärdet. Om t.ex. alla i en mänsklig referensklass väljer att äta mycket socker och inte borsta tänderna, även om det är fullt möjligt att äta sockersnål kost och borsta tänderna regelbundet, så kommer i princip alla individer att ådra sig tänder sämre än standarden. Standarden utesluter alltså inte att specifika typer av sjukdomar är vanliga. Den andra aspekten som gör att DPE-teorin kan hantera vanliga sjukdomar är linjedragningen mellan tillräckligt höga och för låga effektiviteter. DPE-teorin drar inte denna linje statistiskt. Istället ges skiljelinjen av en konstant som beskriver ett förhållande mellan organets specifika dispositionsprofil och standarden. Detta innebär att inte heller gränsdragningen mellan tillräckligt höga och för låga effektiviteter begränsar hur stor del av individerna i en referensklass som kan ha en viss typ av patologiskt tillstånd.

Friska situationsspecifika utföranden av fysiologiska funktioner redogörs för av DPE-teorin genom specifika dispositionsprofiler. I de specifika dispositionsprofilerna relateras alla utföranden till situationer. Om vi tittar på den exemplariska kompletta dispositionsprofilen för en mänsklig referensklass så kommer dess specifika dispositionsprofiler visa att utförandet av många olika fysiologiska funktioner varierar med situationen. Det är t.ex. bättre att magen utsöndrar maximalt med matsmältningssyror endast när det finns mat i matsmältningssystemet än att magen gör detta konstant. Att ha en lägre aktivitet av en viss fysiologisk funktion i en viss situation, i vilken det är lämpligt att ha en lägre aktivitet, blir därför inte patologiskt, utan friskt, enligt DPE-teorin.

Situationsspecifika sjukdomar är inte heller något problem för DPE-teorin. Vad som tas i beaktande i bestämmandet av ett organs hälsostatus är endast hur dess specifika dispositionsprofiler ser ut jämfört med standarden. Det spelar ingen roll i vilken situation individen i fråga befinner sig. Om individen är i en situation som skadat organet så att det inte längre är disponerat att utföra sin typs fysiologiska funktioner lika bra som standarden, så har organets specifika dispositionsprofiler nedsatt effektivitet. Det spelar då ingen roll huruvida motsvarande organ hos andra individer i referensklassen skulle skadas på samma sätt om de hamnade i samma situation. Så länge situationen inte är oundviklig eller vanlig för individer i referensklassen så drabbar inte skadan den exemplariska kompletta dispositionsprofilen. Det innebär att effektiviteten hos det skadade organets specifika dispositionsprofil är nedsatt. Givet att effektiviteten är signifikant nedsatt så är organet patologiskt.

DPE-teorin klarar också att redogöra för patologiska tillstånd som består i att ett organ eller annat fysiologiskt system utför en viss fysiologisk funktion även i en situation där det adekvata vore att inte utföra den. För att ett organ ska räknas som friskt enligt DPE-teorin måste dess specifika dispositionsprofiler ha tillräckligt hög effektivitet. De specifika dispositionsprofilerna anger dock inte enbart hur fysiologiska funktioner utförs i situationer i vilka det är relevant att utföra funktionen i fråga. De rapporterar om utförande i alla situationer, alltså även i situationer i vilka det inte är adekvat att utföra funktionen i fråga.

Sammanfattningsvis klarar DPE-teorin alltså samtliga fyra utmaningar om klarhet, rimliga implikationer, värdefrihet och referensklass-relativism och integration väl. Den klarar dem dessutom bättre än BST.

Låt mig avslutningsvis, för den som är intresserad av att läsa vidare i avhandlingen, berätta hur avhandlingen är disponerad i olika kapitel. Kapitel 1 introducerar den filosofiska debatten om hälsa och patologi och positionerar teorin jag utvecklar i avhandlingen i relation till denna debatt. Här sätter jag också upp ett antal desiderata för teorin som jag utvecklar. Kapitel 2 tillägnas helt Boorses biostatistiska teori om hälsa och patologi, samt invändningar mot denna. I kapitel 3-5 utvecklar jag DPE-teorin. Kapitel 3 presenterar grundläggande antaganden som DPE-teorin bygger på och redogör för vad som kan ses som teorins kärna, nämligen dispositionsprofiler. Kapitel 4 fokuserar på referensklasser och standarder för hälsa. Kapitel 5 tillhandahåller ett effektivitetsmått och använder detta för att definiera ’hälsa’ och ’patologi’. Efter detta diskuterar jag i kapitel 6 och 7 implikationer av DPE-teorin. Kapitel 6 handlar om hur vanliga sjukdomar kan redogöras för som patologiska, och om hur friska situationsspecifika utföranden av fysiologiska funktioner kan redogöras för som friska samtidigt som situationsspecifika sjukdomar redogörs för som patologiska. Kapitel 7 diskuterar vad DPE-teorin implicerar med avseende på olika typer av tillstånd för vilka det kanske inte är intuitivt självklart om de ska räknas som friska eller patologiska. Typerna av tillstånd det handlar om är (i) tillstånd som är triviala för organismen som helhet (t.ex. en enstaka ickefungerande cell), (ii) normalt åldrande (t.ex. ålderssynthet), (iii) försvarsmekanismer (t.ex. feber) och (iv) risktillstånd (t.ex. högt blodtryck). Kapitel 8 avslutar. Det finns också ett appendix där jag samlat DPE-teorins definitioner.

References

Agich, G. J. 1983. Disease and Value: A Rejection of the Value-Neutrality Thesis. *Theoretical Medicine and Bioethics*, 4(1): 27-41.

Amundson, R. 2000. Against Normal Function. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 31(1): 33-53.

Bigelow, J., and Pargetter, R. 1987. Functions. *The Journal of Philosophy*, 84(4): 181-196.

Boorse, C. 1976a. What a Theory of Mental Health Should Be. *Journal for the Theory of Social Behaviour*, 6(1): 61-84.

———. 1976b. Wright on Functions. *Philosophical Review*, 85(1): 70-86.

———. 1977. Health as a Theoretical Concept. *Philosophy of science*, 44(4): 542-573.

———. 1987. Concepts of Health. In: *Health Care Ethics: An Introduction* (pp. 359-393), D. VanDeVeer, and T. Regan (eds). Philadelphia: Temple University Press.

———. 1997. A Rebuttal on Health. In: *What is Disease?* (pp. 1-134), J. M. Humber, and R. F. Almeder (eds). Totowa, NJ: Humana Press.

———. 2002. A Rebuttal on Functions. In: *Functions* (pp. 63-112), A. Ariew, R. Cummins, and M. Perlman (eds). New York: Oxford University Press.

———. 2010. Disability and Medical Theory. In: *Philosophical Reflections on Disability. Philosophy and Medicine, vol. 104* (pp. 55-88), D. Ralston, and J. Ho (eds). Dordrecht: Springer.

———. 2011. Concepts of Health and Disease. In: *Philosophy of Medicine* (pp. 13-64), D. M. Gabbay, F. Gifford, P Thagard, and J. Woods (eds). Oxford: Elsevier.

———. 2012. Clinical Normality. Draft for the conference *Boundaries of Disease*, Hamburg.

———. 2014. A Second Rebuttal on Health. *Journal of Medicine and Philosophy*, 39(6): 683-724.

Broadbent, A. 2017. Health as a Secondary Property. *British Journal for the Philosophy of Science*, 70(2): 609-627.

Brown, W. M. 1985. On Defining 'Disease'. *Journal of Medicine and Philosophy*, 10(4): 311-328.

Brülde, B., and Tengland, P-A. 2003. *Hälsa och sjukdom: en begreppslig utredning*. Lund: Studentlitteratur AB.

Cooper, R. 2005. *Classifying Madness*. Dordrecht: Springer.

Clouser, K. D., Culver, C. M., and Gert, B. 1997. Malady. In: *What is Disease?* (pp. 173-218), J. M. Humber and R. F. Almeder (eds.) Totowa, NJ: Humana.

Culver, C. M. and Gert, B. 1982. *Philosophy in Medicine: Conceptual and Ethical Issues in Medicine and Psychiatry*. Oxford: Oxford University press.

Cummins, R. 1975. Functional Analysis. *The Journal of Philosophy*, 72(20): 741-765.

Dawber, T. R. 1980. *The Framingham Study: The Epidemiology of Atherosclerotic Disease*. Cambridge, MA: Harvard University Press.

DeVito, S. 2000. On the Value-Neutrality of the Concepts of Health and Disease: Unto the Breach Again. *Journal of Medicine and Philosophy*, 25(5): 539-567.
Doust, J., Walker, M. J., and Rogers, W. A. 2017. Current Dilemmas in Defining the Boundaries of Disease. *Journal of Medicine and Philosophy*, 42(4): 350-366.
Dussault, A. C., and Gagné-Julien, A-M. 2015. Health, Homeostasis, and the Situation-Specificity of Normality. *Theoretical Medicine and Bioethics*, 36(1): 61-81.
Engelhardt, H. T. Jr. 1975. The Concepts of Health and Disease. In: *Evaluation and Explanation in Biomedical Sciences* (pp. 125-141), H. T. Engelhardt, JR. and S. F. Spicker (eds.), Dordrecht: D. Reidel.
———. 1976. Ideology and Etiology. *Journal of Medicine and Philosophy*, 1(3): 256-268.
Ereshefsky, M. 2009. Defining 'Health' and 'Disease'. *Studies in History and Philosophy of Biological & Biomedical Sciences*, 40(3): 221-227.
Forest, D., and Le Bidan, M. 2016. In Search of Normal Functions: BST, Cummins Functions, and Hempel's Problem. In: *Naturalism in the Philosophy of Health* (pp. 39-51), É. Giroux (ed). Springer International Publishing Switzerland. (eBook)
Fulford, K. W. M. 1989. *Moral Theory and Medical Practice*. NY: Cambridge University Press.
Garson, J., and Piccinini, G. 2014. Functions Must be Performed at Appropriate Rates in Appropriate Situations. *British Journal for the Philosophy of Science*, 65(1): 1-20.
Giroux, E. 2015. Epidemiology and the Bio-Statistical Theory of Disease: A Challenging Perspective. *Theoretical Medicine and Bioethics*, 36(3): 175-195.
Godfrey-Smith, P. 1993. Functions: Consensus without Unity. *Pacific Philosophical Quartely*, 74(3): 196-208.
Guerrero, J. D. 2010. On a Naturalist Theory of Health: A Critique. *Studies in History and Philosophy of Biological & Biomedical Sciences*, 41(3): 272-278.
Hare, R. M. 1986. Health. *Journal of Medical Ethics*, 12(4): 174-181.
Hausman, D. 2011. Is an Overdose of Paracetamol Bad for One's Health? *British Journal for the Philosophy of Science*, 62(3): 657-668.
———. 2012. Health, Naturalism, and Functional Efficiency. *Philosophy of Science*, 79(4): 519-541.
Hesslow, G. 1993. Do We Need a Concept of Disease? *Theoretical Medicine*, 14(1): 1-14.
Kingma, E. 2007. What is it to be Healthy? *Analysis*, 67(2): 128-133.
———. 2010. Paracetamol, Poison, and Polio: Why Boorse's Account of Function Fails to Distinguish Health and Disease. *British Journal for the Philosophy of Science*, 61(2): 241-261.
———. 2014. Naturalism about Health and Disease: Adding Nuance for Progress. *Journal of Medicine and Philosophy*, 39(6): 590-608.
———. 2016. Situation-Specific Disease and Dispositional Function. *British Journal for the Philosophy of Science*, 67(2): 391-404.
———. 2017. Disease as Scientific and as Value-Laden Concept. In: *Handbook of the Philosophy of Medicine* (pp. 45-63), T. Schramme, and S. Edwards (eds). Dordrecht: Springer Netherlands.
Kitcher, P. 1993. Function and Design. *Midwest Studies in Philosophy*, 18(1): 379-397.

Kraemer, D. M. 2013. Statistical Theories of Functions and the Problem of Epidemic Disease. *Biology and Philosophy*, 28(3): 423-438.
Lemoine, M., and Giroux, É. 2016. Is Boorse's Biostatistical Theory of Health Naturalistic? In: *Naturalism in the Philosophy of Health* (pp. 19-38), É. Giroux (ed). Springer International Publishing Switzerland. (eBook)
Merskey, H. 1986. Variable Meanings for the Definition of Disease. *Journal of Medicine and Philosophy*, 11(3): 215-232.
Millikan, R. G. 1993. *White Queen Psychology and Other Essays for Alice*. Cambrige, MA: MIT Press.
———. 1984. *Language, Thought and Other Biological Categories*. Cambridge, MA: MIT Press.
Neander, K. 1991. Functions as Selected Effects: The Conceptual Analyst's Defense. *Philosophy of Science*, 58(2): 168-184.
Nordenfelt, L. 1987. *On the Nature of Health. An Action-Theoretic Approach*. Dordrecht: Kluwer.
———. 1993. *Quality of Life, Health and Happiness*. Aldershot: Ashgate Publishing.
———. 2001. *Health, Science, and Ordinary Language*. New York: Rodopi.
———. 2018. Functions and Health: Towards a Praxis-Oriented Concept of Health. *Biological Theory*, 13(1): 10-16.
Pawelzik, 1990. Krankheit als Funktionsgestörtheit: Eine Kritik an C. Boorses ‚objektiver' Krankheitstheorie. *Analyse & Kritik*, 12(1): 5-33.
Plunkett, D., and Cappelen, H. 2020. A Guided Tour of Conceptual Engineering and Conceptual Ethics. In: *Conceptual Engineering and Conceptual Ethics* (pp. 1-26), H. Cappelen, D. Plunkett, and A. Burgess (eds). Oxford: Oxford University Press.
Reznek, L. 1987. *The Nature of Disease*. London: Routledge and Kegan Paul.
Regionala cancercentrum i samverkan. 2014. *Akut myolisk leukemi: Nationellt vårdprogram*. (https://cancercentrum.se/globalassets/cancerdiagnoser/blod-lymfom-myelom/aml/natvp_akut_myeloisk_leukemi_maj2014.pdf)
Scadding, J. G. 1988. Health and Disease: What Can Medicine do for Philosophy? *Journal of Medical Ethics*, 14(3): 118-124.
Schaffner, K. F. 1993. *Discovery and Explanation in Biology and Medicine*. Chicago: University of Chicago Press.
Schroeder, A. 2013. Rethinking Health: Health or Healthier Than. *The British Journal for the Philosophy of Science*, 64(1): 151-159.
Schwartz, P. 2007. Defining Dysfunction: Natural Selection, Design, and Drawing a Line. *Philosophy of Science*, 74(3): 364-385.
———. 2008. Risk and Disease. *Perspectives in Biology and Medicine*, 51(3): 320-334.
———. 2014. Reframing the Disease Debate and Defending the Biostatistical Theory. *Journal of Medicine and Philosophy*, 39(6): 572-589.
Simon, J. 2007. Beyond Naturalism and Normativism. *Philosophical Papers*, 36(3): 343-370.
Sorensen, R. 2018. Vagueness. In: *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), Edward N. Zalta (ed). (https://plato.stanford.edu/archives/sum2018/entries/vagueness/)
Stempsey, W. E. 2000. *Disease and Diagnosis: Value-Dependent Realism*. Dordrecht: Springer.
Szasz, T. 1974. *The Myth of Mental Illness*. New York: Harper & Row.

Täljedal, I-B. 2004. Strong Holism, Weak Holism, and Health. *Medicine, Health Care and Philosophy*, 7(2): 143-148.
Vácha, J. 1985. German Constitutional Doctrine in the 1920s and 1930s and Pitfalls of the Contemporary Conception of Normality in Biology and Medicine. *Journal of Medicine and Philosophy*, 10(4): 339-368.
van der Steen, W. J., and Thung, P. J. 1988. *Faces of Medicine: A Philosophical Study*. Dordrecht: Kluwer.
Varga, S. 2020. Epistemic Authority, Philosophical Explication, and the Bio-Statistical Theory of Disease. *Erkenntnis*, 85(4): 937-956.
Wachbroit, R. 1994. Normality as a Biological Concept. *Philosophy of Science*, 61(4): 579-591.
Wakefield, Y. 1992. The Concept of Mental Disorder. On the Boundary between Biological Facts and Social Value. *American Psychologist*, 4(3):373-388.
Werkhoven, S. 2020. Health and Reference Classes. *Journal of Medicine and Philosophy*, 45: 145-158.
Worrall, J. and Worrall, J. 2001. Defining Disease: Much Ado about Nothing? In: *Life Interpretation and the Sense of Illness Within the Human Condition* (pp. 33-55), A-T Tymieniecka, and E. Agazzi (eds). Dordrecht: Kluwer Academic Publishers.
World Health Organization. 1948. *Preamble to the Constitution of WHO as adopted by the International Health Conference, New York, 19 June - 22 July 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of WHO, no. 2, p. 100) and entered into force on 7 April 1948.*
Wright, L. 1973. Functions. *Philosophical Review*, 82(2): 139-168.

ISBN 978-91-7911-524-1

Department of Philosophy

Stockholm University